

```

Db      945 TTCAAGGATTTCTCTCCTCAGCCTCCTAGTATGCTGGATTAACAGGCAACCCGACACAC
Oy      2872 GCCGTGATTAATTTTTTTGATATTTTATAGACAGAGGTTTACCATATGTTGGCAGGCTG
Db      1005 GCCGTGATTAATTTTTTTGATATTTTATAGACAGAGGTTTACCATATGTTGGCAGGCTG
Oy      2932 GTCTCGAATCCCTGACCTTAGGTGATCCACCTGCTGACCTTCCCAAGTGTGGATTA
Db      1065 GTCTCGAATCCCTGACCTTAGGTGATCCACCTGCTGACCTTCCCAAGTGTGGATTA
Oy      2992 TAGCATGAGCAGCAGTGGCCAGTATTAAGAAAGTTAAAGGACATGGCAATGACAC
Db      1125 TAGCATGAGCAGCAGTGGCCAGTATTAAGAAAGTTAAAGGACATGGCAATGACAC
Oy      3052 GCCTATCTAGCTTTCCCTGCCAAAGCAAGGCAAGCTCTGGGCTACCTTTCTGGGT
Db      1185 GCCTATCTAGCTTTCCCTGCCAAAGCAAGGCAAGCTCTGGGCTACCTTTCTGGGT
Oy      3112 TCTACTTCCAAAGGAGCAGTCAAGAACTGGAGGCTTGGAGACCACTTCATCCACTCT
Db      1245 TCTACTTCCAAAGGAGCAGTCAAGAACTGGAGGCTTGGAGACCACTTCATCCACTCT
Oy      3172 AGGCTCCCTATGGAGAGTTAGGTCAGAGCAGGAAAGGTCCTGACAGGCTGACCA
Db      1305 AGGCTCCCTATGGAGAGTTAGGTCAGAGCAGGAAAGGTCCTGACAGGCTGACCA
Oy      3232 GGGCTCTGATCCCTTCAAGAACCCCAATCGGTCTCTCTACAGAGCCCAAGCCAC
Db      1365 GGGCTCTGATCCCTTCAAGAACCCCAATCGGTCTCTCTACAGAGCCCAAGCCAC
Oy      3292 CTGCTGACAGCCAGCTGCTGCGCATGACCATCACTTACAAAGCAAGTGGCTAATGCC
Db      1425 CTGCTGACAGCCAGCTGCTGCGCATGACCATCACTTACAAAGCAAGTGGCTAATGCC
Oy      3352 GCTTAGGCTCTCTCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Db      1485 GCTTAGGCTCTCTCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Oy      3412 ATGGCAGTTCCTTAATCTTCTCTGCTCTGCTACTACTATCAATCCGCTTATTTATAGTAA
Db      1545 ATGGCAGTTCCTTAATCTTCTGCTCTGCTACTACTATCAATCCGCTTATTTATAGTAA
Oy      3472 GCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      1605 GCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Oy      3532 TCCCTGGGCGCTCCAGCAGCTCAGGGCCAGTGCCAGTCCACTTACAACTAAGCT
Db      1658 TCCCTGGGCGCTCCAGCAGCTCAGGGCCAGTGCCAGTCCACTTACAACTAAGCT
Oy      3592 GGGCTCTGACAGCTCTCTGCGCATGAGCTGAGCTGCGCTGCGCTGCGCTGCGCTGCG
Db      1718 GAGCTCTGACAGCTCTCTGCGCATGAGCTGAGCTGCGCTGCGCTGCGCTGCGCTGCG
Oy      3652 TAAAGAGTCACTGAGAGCTGCTCAAGCAGCAGCAGAGGTTTAAAGCACTCTCC
Db      1777 TAAAGAGTCACTGAGAGCTGCTCAAGCAGCAGCAGAGGTTTAAAGCACTCTCC
Oy      3712 TCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1837 TCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Oy      3772 TTTTACTGATCCCAAGAGATCTAAGGCCACATTAATTAATTAATTAATTAATTAAT
Db      1897 TTTTACTGATCCCAAGAGATCTAAGGCCACATTAATTAATTAATTAATTAATTAAT
Oy      3892 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      1957 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Oy      3892 CTCAAGCAATCCCTGCTTACCTCCCAAGGCGTGGATTAACAGTGTGAGCTCTG
Db      2017 CTCAAGCAATCCCTGCTTACCTCCCAAGGCGTGGATTAACAGTGTGAGCTCTG
Oy      3952 CACTTGACCAACACATGCTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Db      2077 CACTTGACCAACACATGCTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Oy      4012 TCCATCACAGGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG
Db      2137 TCCATCACAGGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG
Oy      4072 CAGTGCAGAGATCTCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      2197 CAGTGCAGAGATCTCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Oy      4132 ACCAGGCTGCTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Db      2257 ACCAGGCTGCTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Oy      4192 GTTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      2317 GTTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Oy      4252 AGTCTGGATTTACAGTGTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACG
Db      2377 AGTCTGGATTTACAGTGTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACG
Oy      4312 TATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2437 TATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4372 GCATTTGAGGGGCAAGGTGGGGATCACTGAGCTGAGGATTAAGGTTGGGCAACA
Db      2497 GCATTTGAGGGGCAAGGTGGGGATCACTGAGCTGAGGATTAAGGTTGGGCAACA
Oy      4432 TAGTGAGACCCGCTCTACCAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2557 TAGTGAGACCCGCTCTACCAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4492 TGTGCTCCAGCTACTTGGGAAGTGAAGTGGGGATGCTGTAAGCTGAGAGTGGAG
Db      2617 TGTGCTCCAGCTACTTGGGAAGTGAAGTGGGGATGCTGTAAGCTGAGAGTGGAG
Oy      4552 GCTGCACTGAGCTATGATCAACACACCTGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Db      2677 GCTGCACTGAGCTATGATCAACACACCTGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAG
Oy      4612 GCAACCAAAATATATATATATATATATATATATATATATATATATATATATATATAT
Db      2737 GCAACCAAAATATATATATATATATATATATATATATATATATATATATATATATAT
Oy      4672 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2797 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4732 TGTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2857 TGTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4792 TTTAAAAAAAGTTAGGCGCAGCCACAGGCGCTCAACCTGTAATCCAGCACTTGGGAG
Db      2917 TTTAAAAAAAGTTAGGCGCAGCCACAGGCGCTCAACCTGTAATCCAGCACTTGGGAG
Oy      4852 GCCAAGGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2977 GCCAAGGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4912 CTGATCTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      3037 CTGATCTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Oy      4972 TCCAGCACTTTGGAGGCGGAGGCGGT
Db      3097 TCCAGCACTTTGGAGGCGGAGGCGGT
```


XX WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX Olek A. Piepenbrock C. Berlin K;
 XX WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1; SEQ ID NO 2219; 32pp + Sequence Listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, ananemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/intestinal bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SO Sequence 6396 BP; 1584 A; 89 C; 1657 G; 3066 T; 0 other;

Query Match 25.7%; Score 1284.2; DB 24; Length 6396;
 Best Local Similarity 66.7%; Pred. No. 1.5e-313;
 Matches 2107; Conservative 0; Mismatches 694; Indels 357; Gaps 5;

QY 122 GCACCCACACGACGATACCTATGTAACACACCTGCACCATGATACCTAT 181
 DB 3596 GTAAATTATATGATGATGATGATGATGATGATGATGATGATGATGAT 3654
 QY 182 GTAAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 DB 3655 GTAA-TTAAATTGATGATGATGATGATGATGATGATGATGATGATGAT 3712
 QY 242 GTGTGTAGAAAATCACTGCAATCTCAAGATAGTAAAGCTTAACTTCAAG 301
 DB 3713 GTGTGTAGAAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3772
 QY 302 GAGGTGACAGAAAGGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 361
 DB 3773 GAGGTGATGAAAGGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 3832
 QY 362 CAGGCTCTGGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 421
 DB 3833 TAGGCTTTTGGAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 3892
 QY 422 TTTAAATGCAACATCCCTAAGCCACAGACTCAAGCTGAGAGAAAGTCC 481
 DB 3893 TTTAAATGATATATTTTAAAGTATATGATTTTAAAGTATATTTTAA 3952
 QY 482 CCCGTTTAAATTAATTTTGGGCGATTTCTTACGCGCTTAAAGACCAAC 541
 DB 3953 TTGCTTTTAAATTAATTTTGGGCGATTTTACGCTTTTAAAGATTTAA 4012
 QY 542 CTAGAGCTGCTCTCTTCAATGAAACAATAGAGAGAGTGTAGTAAAGCC 601
 DB 4013 TTAGAGTTTGTGTTTATTTTGAATTAATTAAGAGAGTGTAGTAAAT 4072
 QY 602 CTTCACACAGCTTAAAGAGAGCCCTGATGATGATGATGATGATGATG 661
 DB 4073 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4132

QY 662 CATGCTGATTTTCCACAGCTGCCCGTGGGAGTCTGGGCAATTTCCATATTC 721
 DB 4133 TATGTTGAGATTTTATATATTTGTTGCGGAGATTTGGGATTTATTTATTT 4192
 QY 722 GCGTGGCTGGAAGCCAGCAGCATTAACCTCCAGGCTGTCTGTGTACACCC 781
 DB 4193 GGTGTTGGTGAAGTATATGATTAATTTTAAAGGCTGTCTGTGTAAATTT 4252
 QY 782 ACCCCCTCCACAGCCCGGAGGCTTCTCTTCCATCTCTGTGAGCAACT 841
 DB 4253 ATTTTATTTTATGATTTGCGTGTAGGCTTTTATTTTATTTTGAAGTAT 4312
 QY 842 GGGCCCTGCGACCAATCAAGCAGATATGATATGTCACACAGCCCAATCAG 901
 DB 4313 GGGTTTGTGATGATTAATTAATTAAGTAAATCAATGATGATGATGATG 4372
 QY 902 CTCTGCTGACATATGACAGATTTCTGATTTTACAGGCTGATGAAATTC 961
 DB 4373 TTTTGTGATGATTAATTAATTAAGTAAATTTTATTTTATTTTATTTA 4432
 QY 962 CACCATCTTTTCAATTAAGGCGATGAGGCTGAGAGAGAGCTGAACCTAC 1021
 DB 4433 TATTTATTTTATTTTATTAAGGCTGATGAGGCTGAGAGAGATTTATTT 4492
 QY 1022 CACACACACAGGCTGAGGCGTGAACCAACAGACAGCTGTGACAGCCG 1081
 DB 4493 TATTAATTAATTAAGTGTAAAGTTGGATTTGAATTAATTAATTTGAT 4552
 QY 1082 TTTATTTCTTCCATAGCCACAGGCTGTCAAAAGCCAGGCTGATGAG 1141
 DB 4553 TTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTA 4612
 QY 1142 CTTCTCTGAGATTTCTGCGCACAGATTTGAAGTCAACACCCCTTAACCC 1201
 DB 4613 TTTTCTGAGAGCTTTTGTGATTAAGTGAAGTTGATTAAGTATTTTAT 4672
 QY 1202 TCTCTCTGCAAGCCTCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1261
 DB 4673 TTTTCTGAGGCTTTTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 4732
 QY 1262 TTTAGGCGCATGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1321
 DB 4733 TTTAGGCTTATGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4792
 QY 1322 CTTCCAGAAATTTAGAGGCGCATGAGGCTGAGGCTGAGGCTGAGGCT 1381
 DB 4793 TTTTAAAGAAATTTAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4852
 QY 1382 ATTCAAGAAAGGCTTAAGACAGAGCTCTTTGTGAGGCTGAGGCTGAG 1441
 DB 4853 ATTCAAGAAAGGCTTAAGACAGAGCTCTTTGTGAGGCTGAGGCTGAG 4912
 QY 1442 TGAGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1501
 DB 4913 TGAGGCGGCTTATGATTAATTAAGGCTGAGGCTGAGGCTGAGGCTGAG 4972
 QY 1502 GCCACAGAGTCCAGGAGGCTCCACAGCTGAGGCTGAGGCTGAGGCTGAG 1561
 DB 4973 GTTATAGAGTTTAAAGGCTTATTAATTTTATTTTATTTTATTTTAT 5032
 QY 1562 ACCCAGCTGAAACCCAGCTGTGATTAAGGCTGAGGCTGAGGCTGAGGCT 1621
 DB 5033 ATTTATTTGAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5092
 QY 1622 GAGGCTTGAAGTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1681
 DB 5093 GAGGTTTATTAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 5152
 QY 1682 GCACATGCGCAGGCTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1741
 DB 5153 GTATATGCTAGGCTGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 5212
 QY 1742 GCTTGTGTGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1801

Db 5213 GTTTGGGTCTATTAGG---TATAGTTAGTTAGTGGTGTAGTAAATATATAGGAA 5269
QY 1802 AGAATGAGGGGCGCATCAATCACTGACCAAAATATATATAGAGCTCCCTAAAAAAG 1861
Db 5270 AGAATGAGGGGCGTAAATTA----- 5291
QY 1862 AAGTCTCTTCTTCATAGAGAGAGAGAGGGGTTTCTTATATATATAGG 1921
Db 5292 ----- 5291
QY 1922 AGAGCCGCCCTCAAAAAATAGGAGGAGGAGGAGCCCAAGACCCCGTGTGTGT 1981
Db 5292 ----- 5291
QY 1982 TCCAGGGGAGCTCGAACCTTTAGAGGAGCGTGGAGAACCGGTATTCAGGCTCT 2041
Db 5292 ----- 5291
QY 2042 CGAGAGAAAGAGAGCGGCCGCCAAAAATATCCCTCCGGGCGATAGAAATGTTGCC 2101
Db 5292 ----- 5291
QY 2102 TCTTCAAAAAGATGAGAGAGAGCGGAGTTGTATGTGTATATTTTAAACCCAG 2161
Db 5292 ----- 5291
QY 2162 GTAGNNNNNNNNNTGCTTCAATATTTTATGAGCCCTTACAGAAACACAGAG 2221
Db 5292 -----TGTGTATGATTAATTTTATGAGCTTTTATGAGAAATATAGAG 5338
QY 2222 GAGCTTCATCTGAGAGAGAAACAGGAGAAACAGGAGATATCCGTATATATCA 2281
Db 5339 GAGTTTATTTTGGAGAGAAATAGTGGAATATATTTTAAAGGAAATTTTA 5398
QY 2282 GATGATATAGTCTCTAGAAATATCAAGAGGTGAGAGACAGAGCAGCCGTGG 2341
Db 5399 GTAGGATAGGTTTGTAGAAATATATAGTATAGTATAGATATAGATATAGG 5458
QY 2342 CAGTGGGCTATTTCCAGGTGGATGGTGGGAACCTCTTCAAGGAGACCTGCA 2401
Db 5459 TAGTGGGTTTATTTTATTTAGTGGATGGTGGGAATATTTTAAAGGAAATTTGA 5518
QY 2402 GTGGAGAGAACATGACAGATATCTCAGAGAGAGCTTCCAGGAGAGAGATCAGCAG 2461
Db 5519 GTGGAGAGAAATATATAGTATTTTATAGAGAGTTTATTTAGTATAGATATAG 5578
QY 2462 GTGAGAGGCGCTGAGGACCATTCATTAACATCATTTTGGCATCTTACAGCTAGG 2521
Db 5579 GTGAGAGGTTTGGAGTATTTATTAATATATTTTATTTTATTTTATTTAGT 5638
QY 2522 TTCCATATGGAATGGAATATGTTGTTGAGAGAGGCTGCTCCCTCCCATCTTC 2581
Db 5639 TTTTATATGGAATGGAATATGTTGTTGAGAGGCTGTTGTTTATTTTATTTT 5698
QY 2582 TCACACTAGGCTGTGAGAGAGCTGGAGAGCAAGAGATGGGCTGAGAACCTG 2641
Db 5699 TTAATATAGGCTGTGAGAGAGTTGGAGTTAAGATGAGGTTGAGAAATATG 5758
QY 2642 CTAAGCCAGAGAGCTAGCTAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCT 2701
Db 5759 TTTAGTTAGAGAGTTTACGTTAGTGTGAGATTTTGTGTTATTTGTTTGTAT 5818
QY 2702 TGTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2761
Db 5819 TGTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 5878
QY 2762 CTTAGAGCAATGGCGGATCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 2821
Db 5879 TTGAGAGTATATGCGGATTTATTTATTTATTTATTTATTTATTTATTTATTT 5938
QY 2822 TCTCTCTCTCAGCTCTCTGATGAGTATGAGAGCCGACACAGCCCTGAGTAA 2881

Db 5939 TTTTGTATTTAGTTTGTAGTATGGGATATAGTATTCATATAGTTTGGATA 5998
QY 2882 TTTTGTATTTTGTAGAGAGAGGTTTACCATATTTGGCAGGCTGCTCACT 2941
Db 5999 TTTTGTATTTTGTAGAGAGAGGTTTATTTATTTGTTAGGTTGTTGTAAT 6058
QY 2942 CCGACCTTAGTATCAGCTGCTGAGCTTCCCAAGTGTGGATATAGGATAG 3001
Db 6059 TTTGATTTTAGGATATTTATTTGATTTTAACTGTTGGATATAGATAG 6118
QY 3002 CCAGTCCGCCAGCTGATTTATAGAAATTAAGGACATGCAATGCAACGCTATAC 3061
Db 6119 TTTATGCTTTAGTATTTATTAAGATTAAGATATAGTATATGATATGTTATTTAC 6178
QY 3062 GTCTTCCCTCCCAAGAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3121
Db 6179 GTTTTGTATTAAGTAAAGGATTTTGGGTTATTTTGTGCTTTTATTTTAA 6238
QY 3122 AAAGCAGTCAAGTCTGAGGAGGCTTGGAGACCACTTATCATCCTCTAGGCTCTA 3181
Db 6239 AAAGTATGATTAAGATTTGATGAGGCTTGGAGATTTATTTATTTTATTTAGG 6298
QY 3182 TGGAGAGTTGAGTCCAGAGAGGAGGAGGCTGCTGAGAGCTGCTGAGAGGCTCTGA 3241
Db 6299 TGGAGAGTTGAGGTTTATGAGTATGAGGAGGTTTATGATAGTTGATTAAGG 6358
QY 3242 TCCCTACAAACCCCAATGCTGCTCTTACAGG 3279
Db 6359 TTTTATTAATTTTATATCGGCTTTTATTTATTTAGG 6396

RESULT 7
AAS32727/c
ID AAS32727 standard; DNA; 17792 BP.
AAS32727;
17-DEC-2001 (first entry)

Human genomic DNA for novel endocrine antigen, SEQ ID No 681.

Human; endocrine antigen; ds; cytosolic; antinfertility; antidiabetic;
thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
antisense-therapy; antibody; endocrine disorder; hormone imbalance;
reproductive disorder; endocrine cancer; pancreatic disorder;
diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

Homo sapiens.

WO20015319-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US01335.

31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205415.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.

Query Match	7.58;	Score 375.6;	DB 22;	Length 17792;
Best Local Similarity	64.38;	Pred. No. 7.9e-84;		
Matches 748;	Conservative 0;	Mismatches 350;		

QY	3843	GAGAGATGAGAGCTCACGTGTGTCTCCAGGCTGCTCGAAGCTCTGAAAGCTTGAAGGCTCAAGCAATC	3902
Db	8745	GGGAGACGGGGTTTCACCATGTAAGCCAAATGGTGTGGATCTCTGACTCATG - ATC	8658
QY	3903	CCCTCGCTTAGCCTCCCAAGGGGCTGGATTTACAGGTGGAGCTACTGCAATTGACAA	3962
Db	8657	GGCCTCTCTGGCCTCCCAAGTGTGGATTACAGGATGAGCCACCGCG -	8606
QY	3963	CCACATGGTACTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGGTTTCACTCCATCACCA	4022
Db	8605	CCAGCCTTCTCTTTCTTTCTTTTTTTTTTTTTTTTGAAGATGCTGTGCTGTGCGCCA	8546
QY	4023	GCGTGAAGTCAAGTGGGGCAATCTTGGCTCACTGAACCTCTGCTCCCAAGTGCACG	4082
Db	8545	GGCGGAGTGCAGT - GCGGAGATCTTGCTTACTGCACCTCCACTCCAGGTTCAAG	8487
QY	4083	GATTCTCTCTTACCTCTCTGATGATGCTGGAATTTATAGCACACACACAGCGCTGG	4142
Db	8486	AATTCCTCTGCTCACACCTTGGAGATAGGTGATTTACAGGTGCACACACACACACG	8427
QY	4143	CTAATTTTTTTTTTTTTTTCTGATTTTATTTAGTAGAGAGGAGGTTTATCATGTGGCCAGC	4202
Db	8426	CTAATTTTTT - TGATTTTTTATGTAATTTGGTTTACCTCTGTGGCCAGTC	8376
QY	4203	TGGTCTTGAACCCCTGACTCAAGTGAATCCACCCACTCGGCTCCCAAGTCTGGAGAT	4262
Db	8375	TGGTCTGGAATCTCTGACTCAAGTATTTCTCCACCTCGGCTCCCAAGTCTGGAGAT	8316
QY	4263	TACAGTGTACAGCCACCATGCACAGCCACATGGTACATTTTTTAAATTAATTTTTTAAT	4322
Db	8315	TACAGTGTGATGTACACAGCCTGGCTACTGTCTTAATTTTAAAA -	8269
QY	4323	TAAATGTTTTATCTAAGGCCAGTACAGTACCTCGCTGTATCTCCAGACTTTGAG	4382
Db	8268	-----GAGGCGGGGCGAGTGGCTACAGCCTGTATCTCCAGACTTTGGGA	8223
QY	4383	GGCCAAAGTGGGGGATCACTTGAAGCTGGAGTCAAGCTG - GCAACATAGTGA	4439
Db	8222	GGTCAGATGGGTGATCACAAGTCAAGGATTTGAGACCAAGCTGGCTGATATGTGAA	8163
QY	4440	CCCGCTCTCTACCAAAAATTTAAAAATTTAGCTGGGAGTGTGGCATTTGCTGTGGTCC	8104
Db	8102	ACCGCTCTTAGTAAAA - TACAAAAATTTAGCGGGCGTGTGTGCGACCACTGTAGTCC	4499
QY	4500	CAGCTACTGGGAAGTGAAGTGTGGGGATGGTGAAGCTGTGAAGTCAAGGCTGCAGT	4559
Db	8103	CAGCTACTGAGAGGCTGAGGCGAGGAAATCCCTTGAACTCTGGGAGCGGAGGTGGAGT	8044
QY	4560	GAGCTATGATCACACCACTGCACCTTACGCTG - AGTGCAGGCTATCTCAAAAAGCAAC	4617
Db	8043	GAGCTGAGTTGACACCACTGCACCTCAGCCTGGCAACAAAGAGACCTCTCTCAAAA	7984
QY	4618	AAATATATGTTTTATCAAAAGGTAAGGTAATATCAACAGAAATATATGATGACTTTTAAAT	4677
Db	7983	AAAAAAAATATATATATATATATCAAAAAAATTTACCGCGGATGTGTGACACCGCTATA	7924
QY	4678	TGAAAAAGCAATTATGATTTACATGATGATTTAAATATATCAATTAATATATCTTGCTT	4737
Db	7923	ATCCCAAGTACTCGGAGGCTGAGCGAGGAATCTGTGAACTTACGAGAGTAAAGTTTG	7864
QY	4738	C - TTATATATCTAGCAACAGGACATTTGGTTTTTACTAGGCAACCAAGTACTTTA	4795

ID	Accession	Standard	DNA	Length	BP
Db	7863	CAGTTGCTGAGATCTCACCACTCACTCCAGCTGGGCAACAGAGGATTCATCTC	7800		
Qy	4796	AAAAAGTTAGGGCCACGACACAGGGGCTCACACCTGTATATCCAGCACTTTGGAGGCCA	4855		
Db	7803	AAAAAAAAAAAAAGGCCAGGTCGTGGCTCAGCCCTGTATATCCAGCACTTTGGAGGCCA	7744		
Qy	4856	AGCAGAGAGATGTCACCTTGAGGCCAGGAGTTAGACCTGACCAACATAGGAGATCTCGA	4915		
Db	7743	AGGAGAGTGGATTCGCTGATGTCTCAGAGTTGAGAC - AGCTGGCCCAACATGTGAA	7686		
Qy	4916	TCCTGTCTCTATATAAAATTTGACTAGGCCCTTTGGCTTACACCCGTAATCC	4975		
Db	7685	CCCCGCTCTCTACATTAATAAACTACCGGGGTGTCTGGCTGTACTTTAATCTT	7626		
Qy	4976	AGCATTGTGGAGCGCCAGCGGG 4999			
Db	7625	AGCTTTCAAGAGAGCTGAGGACAG 7602			
RESULT 9					
ID	AAK83109	standard	DNA	2816	BP
XX	AAK83109				
AC	AAK83109				
DT	07-NOV-2001	(first entry)			
XX					
DE	Human	Immune/haematopoietic antigen genomic sequence SEQ ID NO:37921.			
KV	Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;				
XX	Cytostatic; gene therapy; vaccine; metastasis; ds.				
XX	Homo sapiens.				
OS	MO200157182-A2.				
PN	09-AUG-2001.				
PD					
XX	17-JAN-2001; 2001WO-US01354.				
PF					
XX	31-JAN-2000; 2000US-0179065.				
PR	04-FEB-2000; 2000US-0180628.				
PR	24-FEB-2000; 2000US-0184664.				
KV	02-MAR-2000; 2000US-0186350.				
PR	16-MAR-2000; 2000US-0189874.				
PR	17-MAR-2000; 2000US-0190076.				
PR	18-APR-2000; 2000US-0198123.				
PR	19-MAY-2000; 2000US-0205515.				
PR	07-JUN-2000; 2000US-0209467.				
PR	28-JUN-2000; 2000US-0214886.				
PR	30-JUN-2000; 2000US-0215135.				
PR	07-JUL-2000; 2000US-0216647.				
PR	07-JUL-2000; 2000US-0216880.				
PR	11-JUL-2000; 2000US-0217487.				
PR	11-JUL-2000; 2000US-0217496.				
PR	14-JUL-2000; 2000US-0218290.				
PR	26-JUL-2000; 2000US-0220963.				
PR	26-JUL-2000; 2000US-0220964.				
PR	14-AUG-2000; 2000US-0224518.				
PR	14-AUG-2000; 2000US-0224519.				
PR	14-AUG-2000; 2000US-0225213.				
PR	14-AUG-2000; 2000US-0225214.				
PR	14-AUG-2000; 2000US-0225266.				
PR	14-AUG-2000; 2000US-0225267.				
PR	14-AUG-2000; 2000US-0225268.				
PR	14-AUG-2000; 2000US-0225270.				
PR	14-AUG-2000; 2000US-0225447.				
PR	14-AUG-2000; 2000US-0225757.				
PR	14-AUG-2000; 2000US-0225758.				
PR	14-AUG-2000; 2000US-0225759.				
PR	18-AUG-2000; 2000US-0226279.				
PR	22-AUG-2000; 2000US-0226681.				
PR	22-AUG-2000; 2000US-0226686.				


```
Db 1708 ACCTGCTCAGCCTCCCAAGTGTAGATAACAGGCGTGAAG---CACACCGGCGCCGG 1764
Qy 3964 CACAGGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGCTTCACTCCATCACCAG 4023
Db 1765 CCCTGCACAACTTTTTTTTTTTTTTTTTTTTTTTTGTGAGAGGAGTCTGTGCTGTACCTGG 1824
Qy 4024 GCTGAGTACAGTGGGGGCAATCTGGCTCAGTAACTTAACTTGGCTCCAGAGTCAACGG 4083
Db 1825 GCTGAGGAGGAGT-GGCCCGATCTCAGTCACTCACTCACTCCGCTCTGCTGATGATGCC 1883
Qy 4084 ATTCTCTGCTTAACTCTCTGAGTACAGTGAATATAGCAGACACACACAGCCGTCGGC 4143
Db 1884 CTCTCTCTCTCAATCTCCCAAGTACGTGGAGTATAGGCGCCACACACAGCCGCGC 1943
Qy 4144 TAATTTTTTTTTTTTTTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 4203
Db 1944 TAATTTTT-----TGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTT 1994
Qy 4204 GGTCTTAACCCCTGACCTCAAGTATCAACCCACCTCGGCTCCCAAGTGGGAT 4263
Db 1995 GGTCTCATGCGCTGACCTC--GTATCGCGCCGCTCACCTCAACCAAGTGGGAT 2052
Qy 4264 ACAGGTCTCAGCCGACATGCACAGCCGACATGTACTTTTAAATTTTAAATTTTAAAT 4323
Db 2053 ACAGGTCTGAGCCACATGCGCGCCGCAAACTATTTTAAATTTTAAATTTTAAATTTTAA 2112
Qy 4324 AAAATGTTATCTAAGCCAGTACAGTACGCTGCTGATTTTAAATTTTAAATTTTAAAT 4383
Db 2113 ACTAAGACATCAGAGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4383
Qy 4384 GCCAAGTGGGGGATCAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4437
Db 2173 GCGGAGTGGGCTGATCCTGAGTGGGAGTTGAGACAGCTGAGCAACATGAGAG 2232
Qy 4438 GACCCGCTCTACCAAAATTTAAATTTAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 4497
Db 2233 AACCCGCTCTACTGAAAA--TACAAATTTAGCCGGGCTGGGCTGGGCTGGGCTGGGCT 2290
Qy 4498 CCCAGTACTTGGGAAGCTGAGGTGTGGGAGTGTGAAGCTGTGAGGTGAGGCTGCA 4557
Db 2291 CCCAGTACTTGGGAGGCTGAGGAGGAGGAGAAATCGCTTGAACCAAGGAGGAGAAATTC 2350
Qy 4558 GTGAGCTATGATCACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4600
Db 2351 GTGAGCCGAGATCGGCGCATTCGACTCCAGCTGCGCAACAG 2393

RESULT 10
AAK84002/c
ID AAK84002 standard; DNA; 15413 BP.
XX
AC AAK84002;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:38814.
XX
KW Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;
XX
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
XX MO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US01354.
XX
PF 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
```


[illegible]

3481 GCTGGCGGGGGGCTGGAAGATGTGCTGGGGCTGGAGCTGGAGCTCTCTGGGG 3540
4288 TCTGTTCACTCTTTTACTAGTAAATCAGCAATATGATCTCTTCCCATTTCTT 4229
3541 CCTCCAGCAGCTCAGGGCCAGTGCACCTCAGTACCAACTAGTGGGCTCTG 3600
4228 TCTTCTCATATTTCTCTCGAATTTGCAAAATTTATTTAGTAAATGATGATTA 4169
3601 ACCAGCTCTGGGAGCTGAGAGCTGAGGCTGGGCTGGGGCTGGGAGAGTAAG 3660
4168 AATTGCAATTCATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4109
3661 CACACTGAGAGGCTGCTCAGCCAGCCAGCCAGGTTTATGCCACCTCTCTCA 3720
4108 ---ATTATTTATTTAGACAGGGCTCTACTCTATTTGCCAGGCTGGAATG 4052
3721 AGGAGAGCCCTTGAGCCAGGCTTTGTCTGGCCCACTCTACTGCTGTTTACT 3780
4051 ATCTGGGCTCTCAAGAACCTCCACCTCTGGGTTCAAGCATCTCTGCTCAG 3995
3781 TCCACACAGACTCATAGGCCACATGATGATTAATAAAGAGAGAGAGAGAGA 3840
3994 ---CCGAGTAGTGAAGACTACAGATGATGTCATGCTGCTGCTGCTGCTGCT 3937
3841 GAGAGAGATGAGCTCTACTGTTGTCTCCAGGCTGGTCTGCAATCTAGCTCA 3900
3936 AGTAGAGATGAGATTCATGATGTTGGCCAGGCTGCTGCTGCTGCTGCTGCT 3877
3901 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
3876 TCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3817
3961 AAC---CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4018
3816 ATCATCCCATCATTTTAAATCTTTTATTTTATTTTATTTTATTTTATTTT 3757
4019 CCCAGCTGAGAGTGGGGGCAATTTGGCTCAGTACCTCTGCTGCTGCTGCTG 4078
3756 CCCAGCTGAGAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3698
4079 AAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4138
3697 AAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3639
4139 CTGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4198
3698 CTGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3588
4199 AGCTGCTGCTGGAACCCCTGACCTCAAGTATGATCACCACCTGGGCTGCC 4258
3587 AGGGGGCTCTTGAAGTACTGATCAGCTGATCACCCTGAGCTGCCAAAGTGTG 3528
4259 GATTTACAGGTGTCAGCCATGACAGCCACATGCT---ACATTTTAAATTT 4312
3527 AGATTACAGGCTGATGATCAGCCGACCTGACAAATATTTTATTTTATTT 3468
4313 ATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4372
3467 TGTATTAGAGATCAAGAAAGACACTGCTGGGACAGTGGCTGATTTCCCA 3408
4373 CATTGAGGGGCAAGGTCGGGATCACTGAGCTGGGAGTTCA---GCTGGGCA 4430
3407 CACTTTGGAGGCAAGCCAGAGATCACTTGAAGCCAGAGTTCAAGACAGCC 3348
4431 ATAGTAGAGCCCTCTCTACCAAAATTTTAAATTTATTTTATTTTATTT 4490
3347 CAACATGGAAGACTCTCTACTAATAAATTTTAAATTTATTTTATTTTATTT 3288
4491 CTGTGTCCAGCTCTCTGGAAGTGAAGTGTGGGATGCTGAAGCTGAGAGTGA 4550
3287 CTGTGTCCAGCTCTCTGAGGCTGAGGAGAGATCTGACACCTCTGGAGTCAA 3228

4551 GGTGAGTGAAGTATGATCAACACACTGCTTACAGCTGAGTACAGC-----GCT 4602
3227 GGTGAGTGAAGTATGATCAACACACTGCTTACAGCTGAGTACAGC-----GCT 3168
4603 ATCTCAAAAGCAACAAATTAATTTATTTATTTATTTATTTATTTATTTATTT 4662
3167 GTCTCAAAAGCAACAAATTAATTTATTTATTTATTTATTTATTTATTTATTT 3108
4663 GATGACATTTTAAATTTGAAAAAGCATTAATGATTAATGATTTTAAATTTCA 4719
3107 TTAGCAGGTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3051

RESULT 11
ABR83568/c
ID ABR83568 standard; DNA; 201143 BP.
XX ABR83568;
DE 29-AUG-2002 (first entry)
XX Human DNA differentially expressed in granulocytic cells #139.
XX Human; de; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX Crohn's disease; ulcerative colitis; periodontal disease;
XX granulocyte activation; chronic inflammation; allergy.
OS Homo sapiens.
XX
XX MO200226999-A2.
XX
XX 11-APR-2002.
XX
XX 03-OCT-2001; 2001WO-US30821.
XX
XX 03-OCT-2000; 2000US-237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI: 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression
XX of genes associated with granulocyte activation, which serves as
XX diagnostic markers that is useful for monitoring disease states and
XX drug toxicity
XX
XX Claim 1; SEQ ID No 139; 114bp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing
XX the expression level to an expression level in an unactivated
XX GC, where differential expression of Gs is indicative of GCA.
XX Also included are modulating (M2) Gs by contacting GC with an agent
XX that alters the expression of at least one gene in Gs; (2) screening (M3)
XX for an agent capable of modulating GCA or an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease using the
XX gene expression profile; (3) detecting (M4) an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease (especially
XX subject to a pathogen or sterile inflammatory disease, by detecting the
XX level of expression in a sample of the tissue of gene(s) from Gs, where
XX the level of expression of the gene is indicative of inflammation;
XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,
XX an allergic response in a subject, exposure of a subject to a pathogen
XX or sterile inflammatory disease, by contacting a tissue having

DB 42419 CCACTACTTGGAGGCGAGGAGATCACTTGAATCTGGAGGAGAGGTTGCAG 42360
 OY 4559 TGACCTATGATCACACACCTGACCTTGCAGCTGAGTAC 4597
 DB 42339 TGAGCTGAGTTCACACACCTGACCTTGCAGCTGAGTAC 42321

RESULT 12
 AAK89986
 ID AAK89986 standard; DNA: 32224 BP.
 AC AAK89986;
 DT 05-NOV-2001 (first entry)

DE Human digestive system antigen genomic sequence SEQ ID NO: 3562.
 XX
 KM Human: digestive system antigen; gene therapy; cancer; appendicitis;
 KM ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum; ds.
 XX

OS Homo sapiens.
 XX
 PN WO200155314-A2.
 XX
 PD 02-AUG-2001.
 XX

PE 17-JAN-2001; 2001WO-US01324.
 XX

PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 23-AUG-2000; 2000US-0226279.
 PR 23-AUG-2000; 2000US-0226681.
 PR 23-AUG-2000; 2000US-0226686.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236337.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239837.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241321.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.

CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
50

0 other; 4771 T; 2555 G; 2568 C; 3529 A; 0 other;

Every match 7.0%; Score 350.6; DB 22; Length 13819;
 Best Local Similarity 63.6%; Pred. No. 1.4e-77;
 Matches 661; Conservative 0; Mismatches 319; Indels 50.

[illegible]

Oy		4918	TTCGCTCATATAAAAATTAAAAATGGCTAGGCCCTTGGCTTTACACCCTTAATCCCG	49777
Dd		1607	CTGTCTCACAGAAATACA AAAATTACTGGGCATGTGGCGCAGCTGGCAATCCAG	1548B
Oy		4978	CACCTTGGGAGCCGAGGC	4996
Dd		1547	CTACTCGGTA GCTGAAGC	1529

RESULT 14
AAK67239/c
ID AAK67239 standard; DNA; 30393 BP.
XX

AC AAK67239;

DT	06-NOV-2001	(first entry)
XX		

Accession	Antigen	Genomic sequence	Seq ID	NO:22051.
XX				

cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

PN WO200157182-A2.
yy

PD 09-AUG-2001.
YV

17-JAN-2001; 2001WO-US01354

PR 04-FEB-2000: 2000TS-01B0638

02-MAR-2000; 2000US-0186350

17-MAR-2000; 2000US-0190076.

19-MAY-2000; 2000US-0205515.

28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000TS-031513E

07-JUL-2000; 2000US-0216880

11-JUL-2000; 2000US-0217496.

26-JUL-2000; 2000US-0220963.

14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224518.

14-AUG-2000; 200005-0225213.
14-AUG-2000; 200005-0225214

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279

22-AUG-2000; 2000US-0226868.

23-AUG-2000; 2000US-0227009.

01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229287.01-SEP-2000; 2000TS=0229344.
01-SEP-2000; 2000TS=0229345

05-SEP-2000; 2000US-0229513

PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 20-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249269.
 PR 17-NOV-2000; 2000US-0249269.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX

PA (HUMA-) HUMAN GENOME SCT INC.

PI Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -

PS Disclosure; SEQ ID NO 22051; 3071pp + Sequence Listing; English.

CC AA54951 to AA64702 encode the human immune/hematopoietic antigen (I)
 amino acid sequences given in AA682170 to AA691921. (I) have cytotoxic
 activity, and can be used in gene therapy and vaccine production. (I)
 proteins and polynucleotides may be used in the prevention, diagnosis and
 treatment of diseases associated with inappropriate (I) expression. For
 example, they may be used to treat disorders associated with decreased
 expression by rectifying mutations or deletions in a patient's genome
 that affect the activity of (I) by expressing inactive proteins or to
 supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 the nucleic acids into a host cell and culturing the cell to express the
 protein. (I) proteins and polynucleotides may be used to prevent
 CC diagnosis and cancer metastases of hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-related diseases, especially
 CC to AA687694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AA54942 to AA54950 and AA682169
 CC represent sequences used in the exemplification of the present invention.
 XX

SO Sequence 30393 BP; 8271 A; 7315 C; 8076 G; 6731 T; 0 other;

Query Match 7.0%; Score 349; DB 22; Length 30393;
 Best Local Similarity 72.6%; Pred. No. 5.4e-77;
 Matches 592; Conservative 0; Mismatches 150; Indels 73; Gaps 8;

QY 3844 AGAGATGAGTCTACTGTGTCACAGCTGTGCTGCACTCTAGACCTCAACCAATCC 3903
 DB 11537 AGAGATGAGGTTTACATGTTGTCAGGCTGTGCACTCTAGACCTCAACCAATCC 11478
 QY 3904 CCTGCTTACCTCCCAAGGGCTGGATTACAGGTGTGAGTACCTGACCTGACCAAC 3963
 DB 11477 ACCCGCTTGGCTCCCAAGGTGTGAGTACCTGACCTGACCTGACCTGACCTA 11419
 QY 3964 CACATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4023
 DB 11418 -AATTAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 11360
 QY 4024 GCTGAGTGCAGTGGGGCAATCTTGGCTGACCTGTAACCTGCTCCCAAGGTGACAG 4083
 DB 11359 ACTGAGTGCAGTGGGGCAATCTTGGCTGACCTGTAACCTGCTCCCAAGGTGACAG 11301

Run on:

June 8, 2003, 09:30:48 ; Search time 12606.1 Seconds (without algorithm)

11543.117 Million ce

US-09-622-964-1 copy 1

```
Perfect score: 5000
Sequence: 1 ccaaaaattatctcttc
```

Scoring table: IDENTITY_NUC

Searched: 2054640 Reqs. 145514030

Total number of hits satisfying above

Minimum DB section: 4109280

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hbg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_v1:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mn:*
20: em_com:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_v1:*
30: em_hbg_hum:*
31: em_hbg_in:*
32: em_hbg_lthv:*
33: em_hbg_mus:*
34: em_hbg_pln:*
35: em_hbg_rod:*
36: em_hbg_mem:*
37: em_hbg_vrt:*
38: em_sy:*
39: em_hbg_hum:*
40: em_hbg_mus:*
41: em_hbg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4034.2	80.7	112309	9	AC003025	AC003025 Human Chr
2	4034.2	80.7	142092	9	AFI39813	AFI39813 Homo sapi
3	4034.2	80.7	196080	9	AC004228	AC004228 Homo sapi
4	3254.6	65.1	163915	2	AC087451	AC087451 Homo sapi
5	3254.6	65.1	166867	9	AP003733	AP003733 Homo sapi
6	3142	62.8	133633	2	AC084857	AC084857 Homo sapi
7	2379.8	47.6	160169	2	AC051664	AC051664 Homo sapi
8	1291.8	25.8	6336	6	AX347149	AX347149 Sequence
9	1284.2	25.7	6336	6	AX347148	AX347148 Sequence
10	742.2	14.8	160169	2	AC051664	AC051664 Homo sapi
11	728.2	14.6	133663	2	AC084857	AC084857 Homo sapi
12	435.6	8.7	165423	9	CNS07ECV	CNS07ECV Homo sapi
13	431	8.6	125508	9	HSJ309K20	HSJ309K20 Human Chr
14	405.4	8.1	98599	9	HSJ309K20	HSJ309K20 Human DNA
15	405.4	8.1	121502	9	AC007136	AC007136 Homo sapi
16	398.6	8.0	42265	9	AP000845	AP000845 Homo sapi
17	394.8	7.9	142000	9	AC068889	AC068889 Homo sapi
18	394.6	7.9	186335	2	AC016395	AC016395 Homo sapi
19	392	7.8	166869	9	AL353579	AL353579 Human DNA
20	387.6	7.8	134010	9	AL353579	AL353579 Human DNA
21	387.2	7.7	100827	9	AL591583	AL591583 Human DNA
22	385.2	7.7	10863	9	AL592314	AL592314 Human DNA
23	384.8	7.7	11046	9	AC124612	AC124612 Homo sapi
24	384.6	7.7	11848	2	AC090574	AC090574 Homo sapi
25	383	7.7	172969	9	AL355490	AL355490 Human DNA
26	382.2	7.6	190803	9	AC112117	AC112117 Homo sapi
27	382.2	7.6	189406	9	AC05726.2	AC05726.2 Continuation (3 of
28	382	7.6	110000	2	AC023050	AC023050 Homo sapi
29	380.6	7.6	191474	2	AC023157	AC023157 Homo sapi
30	380.6	7.6	210544	2	AC013399	AC013399 Homo sapi
31	380.6	7.6	171812	9	AC079988	AC079988 Homo sapi
32	380.2	7.6	211305	9	AC090512	AC090512 Homo sapi
33	380.2	7.6	165508	9	AC090512	AC090512 Homo sapi
34	380	7.6	173391	9	AL560465	AL560465 Human DNA
35	378.8	7.6	116103	9	AL583963	AL583963 Human DNA
36	378.4	7.6	80146	9	HS1112F9	HS1112F9 Human DNA
37	378.2	7.6	116368	9	AL390063	AL390063 Homo sapi
38	378.2	7.6	213648	9	HS1251D0	HS1251D0 Human DNA
39	378.2	7.6	193570	2	AL515804	AL515804 Human DNA
40	378.2	7.6	122279	9	AL512402	AL512402 Homo sapi
41	377	7.5	223597	2	AC004517	AC004517 Homo sapi
42	376.8	7.5	44062	9	AC004686	AC004686 Homo sapi
43	376.6	7.5	133478	9	AC097359	AC097359 Homo sapi
44	376.6	7.5	133478	9	AC097359	AC097359 Homo sapi
45	375.6	7.5	187127	9	AC097359	AC097359 Homo sapi

ALIGNMENTS

REFERENCE AUTHORS	ORGANISM	LOCUS	DEFINITION	ACCESSION VERSION	KEYWORDS	SOURCE
1 (bases 1 to 113309) Evans, G. A., Athanasiou, M., Buettner, J., Bumeister, R., Card, P., Gessalboc, F., Dunn, J., English, C., Ethridge, S., Gartner, H. R., Gee, V., Gordon, M., Gotway, G.,	Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.	AC003025	Human Chromosome 11p12.2 PAC clone	AC003025	AC003025.1	HTG.
	Homo sapiens.		112309 bp			
	Homo sapiens		DNA			
			linear			
			PRI 23-JUL-1998			
			complete sequence.			

TITLE
JOURNAL
REFERENCE
AUTHORS
Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N.,
Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S.,
Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
HGS Submission
Unpublished
2 (bases 1 to 112309)

TITLE
JOURNAL
Submitted (21-Oct-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA

REFERENCE
AUTHORS
Evans, G.A., Athanasiou, M., Bradbury, P., Brinac, S., Bumeister, R.,
Buetner, J., Bumeister, R., Card, P., Desaillo, F., Dunn, J.,
English, C., Ehrhardt, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,
Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N.,
Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S.,
Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
Direct Submission

TITLE
JOURNAL
Submitted (23-Jun-1998) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA

COMMENT
IMPORTANT: This submission contains the entire insert of clone
PDJ466a11. PDJ466a11 comes from a PAC library constructed at the
Roswell Park Cancer Institute by the Pieter de Jong group. This
clone has been finished according to strict quality criteria and
attempts have been made to resolve all base calling problems such
as compressions and repetitive elements. The expected Phred/Phrap
calculated errors/10kb is 0.18. In addition, this sequence has
been finished such that 99.9% of consensus base calls consist of
either double-stranded coverage or 2 types of labeling chemistry on
one strand.

Further information regarding the map of this region or
annotation of PDJ466a11 can be found at
http://gestec.swmed.edu/chromosol.htm.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
11p12.2 Best's disease region mapped between STS D11S461 and EST
A1NAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA)
MAPPED CLONE OVERLAP: HGS submitted PAC clones PDJ519013 and
PDJ756b9.

FEATURES
SOURCE
Location/Qualifiers

repeat_region
/organism="Homo sapiens"
/db_xref="taxon:9606"
complement(4341..4656)
/rpt_family="Alu"
repeat_region
complement(9880..10120)
/rpt_family="Alu"
repeat_region
10163..10282
/rpt_family="Alu"
repeat_region
complement(11232..12157)
/rpt_family="Alu"
repeat_region
13351..13560
/rpt_family="L1"
repeat_region
complement(15092..15146)
/rpt_family="MIR"
repeat_region
15826..15932
/rpt_family="MIR"
repeat_region
17011..17146
/rpt_family="MIR"
repeat_region
complement(18018..18313)
/rpt_family="Alu"
repeat_region
18493..18741
/rpt_family="Alu"
repeat_region
18820..18954
/rpt_family="Alu"

repeat_region
complement(21374..21661)
/rpt_family="Alu"
repeat_region
22809..23212
/rpt_family="Alu"
repeat_region
23499..23752
/rpt_family="Alu"
repeat_region
24118..24472
/rpt_family="Alu"
repeat_region
25005..25113
/rpt_family="MIR"
repeat_region
31160..31442
/rpt_family="Alu"
repeat_region
complement(32651..32947)
/rpt_family="Alu"
repeat_region
complement(35970..36025)
/rpt_family="MIR"
repeat_region
38958..39145
/rpt_family="Alu"
repeat_region
44878..45480
/rpt_family="Alu"
repeat_region
49769..50071
/rpt_family="Alu"
repeat_region
complement(50172..50310)
/rpt_family="Alu"
repeat_region
50902..50982
/rpt_family="MIR"
repeat_region
51038..51659
/rpt_family="Alu"
repeat_region
54494..54769
/rpt_family="Alu"
repeat_region
55529..55623
/rpt_family="THE1"
repeat_region
complement(55691..55967)
/rpt_family="Alu"
repeat_region
complement(56049..56257)
/rpt_family="Alu"
repeat_region
56721..57100
/rpt_family="THE1"
repeat_region
58402..58680
/rpt_family="Alu"
repeat_region
complement(58748..59345)
/rpt_family="Alu"
repeat_region
59678..59904
/rpt_family="Alu"
repeat_region
61149..61424
/rpt_family="Alu"
repeat_region
complement(61930..62242)
/rpt_family="Alu"
repeat_region
62276..62710
/rpt_family="MLT1"
repeat_region
62915..63118
/rpt_family="Alu"
repeat_region
complement(63349..63455)
/rpt_family="Alu"
repeat_region
64339..64614
/rpt_family="Alu"
repeat_region
complement(64707..64984)
/rpt_family="Alu"
repeat_region
complement(65596..65811)
/rpt_family="Alu"
repeat_region
complement(66500..66753)
/rpt_family="Alu"
repeat_region
67311..67342
/rpt_family="Alu"
repeat_region
67835..68119
/rpt_family="Alu"
repeat_region
68860..69150
/rpt_family="Alu"
repeat_region
69164..69272
/rpt_family="Alu"
repeat_region
69354..70245
/rpt_family="Alu"
repeat_region
70655..70939

Oy 122 GCAACCCACCATGGACACGATATACCTATGTATACCAACCTGCACACTCATGTATACCTAT 181
 Db 78181 GCAAACACCATGGACACGATATACCTATGTATACCAACCTGCACACTCATGTATACCTAT 78237
 Oy 182 GTAACCAACCTGTGATCTTGACACAGTATCCAGAGCTTTAGAGTGAAGAAAAAAGTG 241
 Db 78240 GTAAACAAACCT-GTACATTTGACACAGTATCCAGAAC-TTAAAGTGAAGAAAAAAGTG 78297
 Oy 242 GTGTGTAGAAAAACCTGCATCTCAGATAGTTAAAGCTTAGTACATTTACAGAGAGA 301
 Db 78298 GTGTGTAGAAAAACCTGCATCTCAGATAGTTAAAGCTTAGTACATTTACAGAGAGA 78357
 Oy 302 GAGGTGCACAGAAAGGAGATGAGTGGGTTTAAAGACAAAGTCTATTTATTAAT 361
 Db 78358 GAGGTGCACAGAAAGGAGATGAGTGGGTTTAAAGACAAAGTCTATTTATTAAT 78417
 Oy 362 CAGGCGTTCTGGAAAGTTAGTCCCAAAACACACATCTATTAATCCCTGCAGTCTTGA 421
 Db 78418 CAGGCGTTCTGGAAAGTTAGTCCCAAAACACACATCTATTAATCCCTGCAGTCTTGA 78477
 Oy 422 TTAATAATGCATCTCCTTAAGGACACAGACTCAGACTGGAGAAAGATCCAGAAAACCTG 481
 Db 78478 TTAATAATGCATCTCCTTAAGGACACAGACTCAGACTGGAGAAAGATCCAGAAAACCTG 78537
 Oy 482 CCGGTTAATAAATTTGGGCGATTTCTTAGGGCTCTAAAGACCAAGAACCACTGGTG 541
 Db 78538 CCGGTTAATAAATTTGGGCGATTTCTTAGGGCTCTAAAGACCAAGAACCACTGGTG 78597

QY	542	CTAGAGCTCTGCTCTCTTCATTTGAACAAATACAAAGAGGTGTGTAAGTAGACACCCACCA	601
Db		78598 CTAGAGCTCTGCTCTCTTCATTTGAACAAATACAAAGAGGTGTGTAAGTAGACACCCACCA	
QY		602 CTTCACAAGACTTAGAGAGCCCTTGACTATGATTTGATTTAAATTTATTTGAATCA	661
Db		78658 CTTCACAAGACTTAGAGAGCCCTTGACTATGATTTGATTTAAATTTATTTGAATCA	
QY		662 CATGCTGAGATTTTACCAGGCTGGCCGGGGGATCTGGCAATTTATTCCTCATTTGGCACT	721
Db		78718 CATGCTGAGATTTTACCAGGCTGGCCGGGGGATCTGGCAATTTATTCCTCATTTGGCACT	
QY		722 GGCTGGCTGGAAGGACGACGATTAATCTCAGGCGTTCTGTGCACACCCGACGACTC	781
Db		78778 GGCTGGCTGGAAGGACGACGATTAATCTCAGGCGTTCTGTGCACACCCGACGACTC	
QY		782 ACCCCCTTCACACAGCCCCGGAGGCTTCTCCTTCATCTCTCTGAACCAATTAAGTAT	841
Db		78838 ACCCCCTTCACACAGCCCCGGAGGCTTCTCCTTCATCTCTCTGAACCAATTAAGTAT	
QY		842 GGAGCCTGCGCAGGCATTCACAGCCAGATTAACGATGATGTACACAGAGCAATTCAGAG	901
Db		78898 GGAGCCTGCGCAGGCATTCACAGCCAGATTAACGATGATGTACACAGAGCAATTCAGAG	
QY		902 CTCTCTGTCAGCATATGACGAATTTCTGATTTTACTAGGCTGATGAATTTCCAGCA	961
Db		78958 CTCTCTGTCAGCATATGACGAATTTCTGATTTTACTAGGCTGATGAATTTCCAGCA	
QY		962 CACCATCTCTTTCAATTAAGGCACTAGGCTGAGAGAGAGCAAGCACTACCGGGGT	1021
Db		79018 CACCATCTCTTTCAATTAAGGCACTAGGCTGAGAGAGAGCAAGCACTACCGGGGT	
QY		1022 CACCAACACAGGTGGCAAGGCTTGGACCAAGAAACAGAGCTTTTACAGCAAGCCGGTA	1081
Db		79078 CACCAACACAGGTGGCAAGGCTTGGACCAAGAAACAGAGCTTTTACAGCAAGCCGGTA	
QY		1082 TTCATTCTTTTCATAGGCCACAGGGCTGTCTAAAGACCCAGGGCCATGACAGAGCTCT	1141
Db		79138 TTCATTCTTTTCATAGGCCACAGGGCTGTCTAAAGACCCAGGGCCATGACAGAGCTCT	
QY		1142 CCTTCTCTGAGAGTTCCTGTGACAGAAAGTTGAAGCTCAGACAGCCCTTAACCCCAAC	1201
Db		79198 CCTTCTCTGAGAGTTCCTGTGACAGAAAGTTGAAGCTCAGACAGCCCTTAACCCCAAC	
QY		1202 TCTCTGCAAGGCCCAAGGCTCAAGAACTGTGTGAGAGATCTCTTTAGCCCTTGTGAT	1261
Db		79258 TCTCTGCAAGGCCCAAGGCTCAAGAACTGTGTGAGAGATCTCTTTAGCCCTTGTGAT	
QY		1262 TTTAGGCGCATGTAGAGAGGGGTGTGGCTTAATTTCAAGCCCTGTCTCAGCCCAACAC	1321
Db		79318 TTTAGGCGCATGTAGAGAGGGGTGTGGCTTAATTTCAAGCCCTGTCTCAGCCCAACAC	
QY		1322 CCTCCAAAGAAATTAAGGGCCATGGCCAGGCTGTCTACCCGTCCTTCTGACGAC	1381
Db		79378 CCTCCAAAGAAATTAAGGGCCATGGCCAGGCTGTCTACCCGTCCTTCTGACGAC	
QY		1382 ATTACAAAGAGACTAAGACAAAGATCTCTTTGTGAGAGTCTGTGCTTAGGAGTCAAG	1441
Db		79438 ATTACAAAGAGACTAAGACAAAGATCTCTTTGTGAGAGTCTGTGCTTAGGAGTCAAG	
QY		1442 TGAGGGGGCTCAGACATCTCAGTGGGCAAGTCCAGGCTCTTAAGATGTGGCAGGGGCACTG	1501
Db		79498 TGAGGGGGCTCAGACATCTCAGTGGGCAAGTCCAGGCTCTTAAGATGTGGCAGGGGCACTG	
QY		1502 GCCACAAGATGCCAGGAGTCCACAGGCTGTGTGCCAGACCTTGTGTGGATCATCGG	1561
Db		79558 GCCACAAGATGCCAGGAGTCCACAGGCTGTGTGCCAGACCTTGTGTGGATCATCGG	
QY		1562 ACCCACTGTGAACCCCACTGTGATTAACAGGTGCCAGGTGTGACTGGGCTGGGCTTT	1621
Db		79618 ACCCACTGTGAACCCCACTGTGATTAACAGGTGCCAGGTGTGACTGGGCTGGGCTTT	

[illegible][illegible]


```
repeat_region /rpt_family="SVA" complement(5503, .6111)
repeat_region /rpt_family="Alu" 5569, .6095
repeat_region /rpt_family="SVA" complement(7239, .7542)
repeat_region /rpt_family="Alu" 7306, .7497
repeat_region /rpt_family="SVA" 7732, .7828
repeat_region /rpt_family="MIR" 9099, .9403
repeat_region /rpt_family="Alu" complement(9526, .9897)
repeat_region /rpt_family="MER1" 9635, .9752
repeat_region /rpt_family="MER1B" 11924, .12197
repeat_region /rpt_family="Alu" complement(11955, .12151)
repeat_region /rpt_family="SVA" complement(12767, .13466)
repeat_region /rpt_family="Alu" 16757, .17039
repeat_region /rpt_family="Alu" complement(16775, .17010)
repeat_region /rpt_family="SVA" complement(18220, .18507)
repeat_region /rpt_family="Alu" 18229, .18485
repeat_region /rpt_family="SVA" 20101, .23446
gene /gene="HUMFERRITIN"
mRNA /note="Human ferritin heavy chain gene" join(20101, .20421, 22218, .22364, 22621, .22746, 22842, .23446)
repeat_region /gene="HUMFERRITIN"
repeat_region /product="ferritin heavy chain" complement(21466, .21731)
repeat_region /rpt_family="Alu" 21506, .21714
gene /rpt_family="SVA" complement(23272, .37375)
gene /gene="VMD2"
mRNA /note="Best macular dystrophy gene" complement(join(23272, .23629, 24840, .25478, 27690, .27841, 28155, .28235, 29427, .29579, 30274, .30351, 30740, .30894, 31787, .32020, 32537, .32631, 35777, .35964, 37308, .37375))
repeat_region /product="VMD2"
repeat_region /product="Best macular dystrophy protein" complement(23735, .23954)
repeat_region /rpt_family="Alu" 23789, .23956
repeat_region /rpt_family="SVA" 26429, .26707
repeat_region /rpt_family="Alu" complement(26444, .26696)
repeat_region /rpt_family="SVA" 27103, .27349
repeat_region /rpt_family="Alu" complement(27106, .27323)
repeat_region /rpt_family="SVA" complement(29161, .29307)
repeat_region /rpt_family="Alu" complement(29762, .30043)
repeat_region /rpt_family="Alu" 29786, .30017
repeat_region /rpt_family="SVA" 30462, .30516
repeat_region /rpt_family="MIR" complement(31095, .31409)
repeat_region /rpt_family="Alu" 33005, .33260
repeat_region /rpt_family="Alu"
```

```
repeat_region 33247, .33597 /rpt_family="L1MC2"
repeat_region complement(33849, .34431) /rpt_family="Alu"
repeat_region 34076, .34411 /rpt_family="SVA"
repeat_region complement(34640, .34890) /rpt_family="Alu"
repeat_region 34972, .35401 /rpt_family="Alu"
repeat_region complement(34988, .35392) /rpt_family="SVA"
repeat_region 36238, .36487 /rpt_family="Alu"
repeat_region complement(36253, .36457) /rpt_family="SVA"
repeat_region complement(38661, .38855) /rpt_family="L1"
repeat_region complement(40997, .41224) /rpt_family="Alu"
repeat_region 41056, .41213 /rpt_family="SVA"
repeat_region 42745, .43020 /rpt_family="Alu"
repeat_region complement(42763, .42966) /rpt_family="SVA"
repeat_region 43312, .43457 /rpt_family="Alu"
repeat_region complement(43963, .44217) /rpt_family="L1"
repeat_region complement(45296, .45533) /rpt_family="SVA"
repeat_region complement(45300, .45401) /rpt_family="L1"
repeat_region complement(45407, .45533) /rpt_family="L1"
repeat_region complement(45542, .45818) /rpt_family="SVA"
repeat_region 45598, .45770 /rpt_family="Alu"
repeat_region complement(45862, .46004) /rpt_family="L1"
repeat_region complement(45998, .46282) /rpt_family="Alu"
repeat_region 46022, .46260 /rpt_family="SVA"
repeat_region complement(46692, .47583) /rpt_family="Alu"
repeat_region 46885, .47215 /rpt_family="SVA"
repeat_region 47363, .47571 /rpt_family="SVA"
repeat_region complement(47665, .47773) /rpt_family="Alu"
repeat_region complement(47792, .48077) /rpt_family="Alu"
repeat_region complement(48110, .48388) /rpt_family="L1"
repeat_region complement(48827, .49102) /rpt_family="Alu"
repeat_region 48883, .49110 /rpt_family="SVA"
repeat_region complement(49395, .49626) /rpt_family="Alu"
repeat_region 50184, .50437 /rpt_family="Alu"
```

```
Query Match 80.7% Score 4034.2 DB 9: Length 142092;
Best Local Similarity 91.9% Pred. No. 0;
Matches 4486; Conservative 0; Mismatches 28; Indels 365; Gaps 7;
QY 122 GCAACCCACACGACACGATACCTATGTAAACCAACCTGCACCATCATGTATACCAT 181
|||||
```

Dp 38764 GCAAAACCACCATGGCAGACGTAATACCTATGTAAACAACCTGCA-CATCATGTATACCTT 38706

Oy 182 GTAAACAACCTGGTATCATCTCTGCACACGATATCCAGAGACTTTAGATGTAAAAAAGTG 241

Dp 38705 GTAAACAACCT-GTACATCTCTGCACACGATATCCAGAAC-TTAAAGTAAAAAAGTG 38648

Oy 242 GTGTAGAAAAAATCACCTGCATATCTCAAGCATAGTTAAAGCTTATGTACATTTAGAGAG 38648

Dp 38647 GTGTAGAAAAAATCACCTGCATATCTCAAGCATAGTTAAAGCTTATGTACATTTAGAGAG 38648

Oy 302 GAGGTGACAGAAAAAGGAGATGAGATGGGTTTAAAGCAAGCTCATTTTAAAT 361

Dp 38587 GAGGTGACAGAAAAAGGAGATGAGATGGGTTTAAAGCAAGCTCATTTTAAAT 38528

Oy 362 CAGGGCTTTCGAAATTAGTCCCAAAACACACATCTCAATATCCCTCAGTGGTGA 421

Dp 38527 CAGGGCTTTCGAAATTAGTCCCAAAACACACATCTCAATATCCCTCAGTGGTGA 421

Oy 422 TTAATATGCAATCCCTTAAGGCACAGACTCAGACTCGAGAAATCCAGAAACTG 481

Dp 38467 TTAATATGCAATCCCTTAAGGCACAGACTCAGACTCGAGAAATCCAGAAACTG 481

Oy 482 CCGGTTTAAATTAACATTTGGGGATCTTAGAGCTCTAAAGACCAAGAACACGCTGC 541

Dp 38407 CCGGTTTAAATTAACATTTGGGGATCTTAGAGCTCTAAAGACCAAGAACACGCTGC 541

Oy 542 CTAGAGCTCTGCTCTCTTCAATGAACAATCAAGAGAGTGTATAGTAGACACCAACCA 38348

Dp 38347 CTAGAGCTCTGCTCTCTTCAATGAACAATCAAGAGAGTGTATAGTAGACACCAACCA 38348

Oy 602 CTTCACACAGCTTTGGAGAGAGCTTGAATGATGATGATTAATTAATTAATGAATCA 38288

Dp 38287 CTTCACACAGCTTTGGAGAGAGCTTGAATGATGATGATTAATTAATTAATGAATCA 38288

Oy 662 CATCTCAGATTTTACACAGCTGCGCGTGGGGATCTGGCATTTATCCATTTTCACCT 38228

Dp 38227 CATCTCAGATTTTACACAGCTGCGCGTGGGGATCTGGGGATTTATCCATTTTCACCT 38228

Oy 722 GGTGGGTGGAACCCAGACAGCATAACTCCAGAGGCTGTCTGCACACCCACACAGACTC 38168

Dp 38167 GGTGGGTGGAACCCAGACAGCATAACTCCAGAGGCTGTCTGCACACCCACACAGACTC 38168

Oy 782 ACCCCCTCTCACAGCCCGCGGAGGTTCTCTTCATCTCTGTAAGCACTTACTGT 841

Dp 38107 ACCCCCTCTCACAGCCCGCGGAGGTTCTCTTCATCTCTGTAAGCACTTACTGT 841

Oy 842 GGGCCCTCGACACCAATCAACAGCCAGATTAACGTATGTCACACAGCCCAATCAGAG 38048

Dp 38047 GGGCCCTCGACACCAATCAACAGCCAGATTAACGTATGTCACACAGCCCAATCAGAG 38048

Oy 902 CTCTCGTCACATATGAGAAATCTGTCAATTTACTAGAGGTGTAATTTCCCAACAA 961

Dp 37987 CTCTCGTCACATATGAGAAATCTGTCAATTTACTAGAGGTGTAATTTCCCAACAA 961

Oy 962 CACATCTTTTCAATAGAGGCACTAGAGCTGAGAGAGAGCTGAATCCTACCCGGGT 1021

Dp 37927 CACATCTTTTCAATAGAGGCACTAGAGCTGAGAGAGAGCTGAATCCTACCCGGGT 1021

Oy 1022 CACACACACAGGTGGAGAGGCTGGACCAAGAAACAGAGACTTTTACTGAGCCGGTA 1081

Dp 37867 CACACACACAGGTGGAGAGGCTGGACCAAGAAACAGAGACTTTTACTGAGCCGGTA 1081

Oy 1082 TTCAATCTTTCCATAGCCACAGAGGCTGTCAAAAGACCCAGGCTTATGTACAGAGCTCT 37808

Dp 37807 TTCAATCTTTCCATAGCCACAGAGGCTGTCAAAAGACCCAGGCTTATGTACAGAGCTCT 37808

Oy 1142 CTTTCTGAGAGTCTCTGCGCAGAGATTTGAAGCTTCAGACAGGCCCCCTAACCCCAAC 37748

Dp 37747 CTTTCTGAGAGTCTCTGCGCAGAGATTTGAAGCTTCAGACAGGCCCCCTAACCCCAAC 37748

Oy 1202 TCTCTCTCAAGGCTCAGAGGCTCAGAAACAATGTTGAGCTCAGACAGGCCCCCTAACCCCAAC 37688

Dp 37687 TCTCTCTCAAGGCTCAGAGGCTCAGAAACAATGTTGAGCTCAGACAGGCCCCCTAACCCCAAC 37688

Dp 37687 TCTCTCTCAAGGCTCAGAGGCTCAGAAACAATGTTGAGCTCAGACAGGCCCCCTAACCCCAAC 37688

QY	1262	TTTTAGGGCCATGTTAGAGGGGGGTTGGCCCTAAATTCAGACCCTGGTCTTCAGCCCAACAC	1321
Db	37627	TTTTAGGGCCATGTTAGAGGGGGGTTGGCCCTAAATTCAGACCCTGGTCTTCAGCCCAACAC	37566
QY	1322	CTCCAGAGAAATTAGAGGGGCCATGCGCAGGGCTGTCTTAGCCGTTCTCTGACGAG	1381
Db	37567	CTCCAGAGAAATTAGAGGGGCCATGCGCAGGGCTGTCTCTGACGAG	37508
QY	1382	ATTACAAAGAGGACTTAGACAAAGACATCTTTGTGAGGTCTGGCTTAGAGATCAG	1441
Db	37507	ATTACAAAGAGGACTTAGACAAAGACATCTTTGTGAGGTCTGGCTTAGAGATCAG	37448
QY	1442	TGACGGGGGCTCAGACATCTACGTGGGCAAGTCCAGCCCTTAAGAGTGGGCAAGGGCACTG	1501
Db	37447	TGACGGGGGCTCAGACATCTACGTGGGCAAGTCCAGCCCTTAAGAGTGGGCAAGGGCACTG	37388
QY	1502	GCCACAGATGCCAGGGAGTCCCAACCGCTAGTCCGCAACCTTCTGTGGGATCATCG	1561
Db	37387	GCCACAGATGCCAGGGAGTCCCAACCGCTAGTCCGCAACCTTCTGTGGGATCATCG	37328
QY	1562	ACCACACTGGAAACCCACCTGTGTAGTACAAAGGTGCCCAAGTGGACTGGGGGCTTT	1621
Db	37327	ACCACACTGGAAACCCACCTGTGTAGTACAAAGGTGCCCAAGTGGACTGGGGGCTTT	37268
QY	1622	GAGGCTTTCAGGGTTGGATGGCCATCTTCCGTAATTTGTGTGGATATGCACACAGCA	1681
Db	37267	GAGGCTTTCAGGGTTGGATGGCCATCTTCCGTAATTTGTGTGGATATGCACACAGCA	37208
QY	1682	GCACATGGGCAAGGTGTGTGGGCACTGTGTGTCTGTGCAAAATGCCCTGAGGGGAAATGA	1741
Db	37207	GCACATGGGCAAGGTGTGTGGGCACTGTGTGTCTGTGCAAAATGCCCTGAGGGGAAATGA	37148
QY	1742	GCTTGTGTGCATCAGAGGCGACGACCGCAGTGTGGCTGCACAAAACACAGGGA	1801
Db	37147	GCTTGTGTGCATCAGAGGCGACGACCGCAGTGTGGCTGCACAAAACACAGGGA	37091
QY	1802	AGAATGAGGGGGGATCAATCATCTGACAAAATTTATATAGAGCTCCCCCTAAAAAAG	1861
Db	37090	AGAATGAGGGGGGATCAATCATCTGACAAAATTTATATAGAGCTCCCCCTAAAAAAG	37068
QY	1862	AAGGTCTCTTTCGATAGAAAGAGGAGAGAGGGGTTTGTCTTATAAATATAAGG	1921
Db	37067	AAGGTCTCTTTCGATAGAAAGAGGAGAGAGGGGTTTGTCTTATAAATATAAGG	37068
QY	1922	AGGAGCCGCCCTCAAAAAATAGGAGGAGGAGACCACAAGCCCGTGGGTGTGTCT	1981
Db	37067	AGGAGCCGCCCTCAAAAAATAGGAGGAGGAGACCACAAGCCCGTGGGTGTGTCT	37068
QY	1982	TCCAGGGGAGCTCGAACCTTTTAGAGGAGCGTGGAGAACCGCTGTAATTCAGGCTCT	2041
Db	37067	TCCAGGGGAGCTCGAACCTTTTAGAGGAGCGTGGAGAACCGCTGTAATTCAGGCTCT	37068
QY	2042	CGAGAGAAAGAGAGCGGCGCCCAAAAAATATCCCTCCCGGGCGATAAAGAAATGTGGCC	2101
Db	37067	CGAGAGAAAGAGAGCGGCGCCCAAAAAATATCCCTCCCGGGCGATAAAGAAATGTGGCC	37068
QY	2102	TCCTCTCAAAAAGATGAGAGAGAACCGGAGTTGTATGTGTAATTTTAAAACTCAG	2161
Db	37067	TCCTCTCAAAAAGATGAGAGAGAACCGGAGTTGTATGTGTAATTTTAAAACTCAG	37068
QY	2162	GTAGNNNNNNNNNNNGCTTCAGTAAATTTTATTTAGGCGCTTCTACGAAACACAAGAG	2221
Db	37067	GTCTCTCAAGTAAATTTTATTTAGGCGCTTCTACGAAACACAAGAG	37022
QY	2222	GAGCTTCCATTTGTGAGAGAAACAGGCGGAAACAGGAGATATCTGTATATTTTCAA	2281
Db	37021	GAGCTTCCATTTGTGAGAGAAACAGGCGGAAACAGGAGATATCTGTATATTTTCAA	36962
QY	2282	GTAGGATTAAGGCTCTAGAAATATATACCAAGGTAGAGAGACACAGACACCGGTGG	2341
Db	36961	GTAGGATTAAGGCTCTCTAGAAATATATCAACCAAGGTAGAGAGACACAGACACCGGTGG	36902

[illegible]

D	35821	CTTAATCTTCTGCTCTGCTACTACATCAATCCGTTATTATTAAGTAAAGCTGCACAGG	35762
Q	3482	CTGGGCGGGGGGCTGGGAAGATGTGCTGGGCTGGGACTGGAGCTCTGGGGG	3541
D	35761	CTGGGCGGGGGGCTGGGAAGATGTGCTGG-----GCTGGAGCTCTGGGGG	35709
Q	3542	CTCCAGCGGCTCAGAGGCGCCAGTGCACCAATCTACATACATTAAGCTGGGCTCTGA	3601
D	35708	CTCCAGCGGCTCAGAGGCGCCAGTGTACCAATCTACATTAAGCTGGGCTCTGA	35649
Q	3602	CCAGCTCTGGGCTCAGAGCTGAGGCTGGGCGCTGGGGGCTGGGCAAGTAAAGATC	3661
D	35648	CCAGGTCCTGGGCTCAGAGCTGAGGCTGGGCGCTGGGGGCTGGGCAAGTAAAGATC	35590
Q	3662	ACACTGAGAGGCTGCTCAAGCCAGGCCAGAGGGTTTTAGCCACCTTCTCTCAACCCA	3721
D	35589	AGACTGAGAGGCTGCTCAAGCCAGGCCAGAGGGTTTTAGCCACCTTCTCTCAACCCA	35530
Q	3722	GGAGGAGCCCTGGAGCCAGGCTTGTGCTGGCCCACTCTACTGGCCCTTTTACTGAAT	3781
D	35529	GGAGGAGCCCTGGAGCCAGGCTTGTGCTGGCCCACTCTACTGGCCCTTTTACTGAAT	35470
Q	3782	CCCACACAGACTCATAGGCCACATAGTACTTAAAAAAGAGAGAGAGAGAGAGAG	3841
D	35469	CCCACACAGACTCATAGGCCACATAGTACTTAAAAAAGAGAGAGAGAGAGAGAG	35410
Q	3842	AGAGAGATGGAGTCTCACTGTGTGTCCAGGCTGCTCCGAATCTTAAGCTCAAGCAAT	3901
D	35409	AGAGAGATGGAGTCTCACTGTGTGTCCAGGCTGCTCCGAATCTTAAGCTCAAGCAAT	35350
Q	3902	CCCCCTGCTTATAGCTCTCCAAAGGGGCTGGATTACAGGTGAGCTACTGACTTGACA	3961
D	35349	CCCCCTGCTTATAGCTCTCCAAAGGGGCTGGATTACAGGTGAGCTACTGACTTGACA	35290
Q	3962	ACCAATAGTACTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGTTTCACTCATACCC	4021
D	35289	ACCAATAGTACTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGTTTCACTCATACCC	35230
Q	4022	AGGCTGAGTGAAGTGGGGCAATCTTGCTCACTGTACCTGCTCCAGGTGCAAG	4081
D	35229	AGGCTGAGTGAAGTGGGGCAATCTTGCTCACTGTACCTGCTCCAGGTGCAAG	35170
Q	4082	CGAATTCGCCCTTATAGCTCTCCAGGTGAGTGTGAATTAAGCACACACCAACGCTG	4141
D	35169	CGAATTCGCCCTTATAGCTCTCCAGGTGAGTGTGAATTAAGCACACACCAACGCTG	35110
Q	4142	GCTAATTTTTTTTTTTTTTTTCTGATTTTAAAGTAAAGAGAGGTTTCAATGTTGGCAG	4201
D	35109	GCTAATTTTTTTTTTTTTTTTCTGATTTTAAAGTAAAGAGAGGTTTCAATGTTGGCAG	35050
Q	4202	CTGGTCTTGAACCCCTGACCTCAAGTATCCACCCACCTCGGCTCCCAAAGTGTGGGA	4261
D	35049	CTGGTCTTGAACCCCTGACCTCAAGTATCCACCCACCTCGGCTCCCAAAGTGTGGGA	34990
Q	4262	TTACAGGTGAGCCACATGACAGCCCAACATGATTTTAAATTTATTTTAA	4321
D	34989	TTACAGGTGAGCCACATGACAGCCCAACATGATTTTAAATTTATTTTAA	34930
Q	4322	TTAAATTTTATCTAAGGCCAGTGAAGTGTGCGCTGTTAATCCAGACATTTGAG	4381
D	34929	TTAAATTTTATCTAAGGCCAGTGAAGTGTGCGCTGTTAATCCAGACATTTGAG	34870
Q	4382	GGGCGCAAGGTGGGGGATCACTGAGCTGGGAGTTTCAAGCTGGGCAACATAGTAGACC	4441
D	34869	GGGCGCAAGGTGGGGGATCACTGAGCTGGGAGTTTCAAGCTGGGCAACATAGTAGACC	34810
Q	4442	CCGTCTACACAAAAATTTAAAAATTAAGCTGGGAGTGTGGCATTTCCCTGGTCCCA	4501
D	34809	CCGTCTACACAAAAATTTAAAAATTAAGCTGGGAGTGTGGCATTTCCCTGGTCCCA	34750
Q	4502	GCTACTTGGAAAGTGAAGTGTGGGATGAGCTGGAAGCCTGTGAGTCAAGCTCAAGTA	4561


```
repeat_region complement(52700, .52979)
/rpt_family="Alu"
repeat_region 53281, .53568
/rpt_family="Alu"
repeat_region 54071, .54204
/rpt_family="Alu"
repeat_region 54506, .54793
/rpt_family="Alu"
repeat_region 54807, .55061
/rpt_family="Alu"
repeat_region complement(64533, .64826)
/rpt_family="Alu"
repeat_region complement(65245, .65783)
/rpt_family="Alu"
repeat_region complement(68465, .68732)
/rpt_family="Alu"
repeat_region complement(70191, .70476)
/rpt_family="Alu"
repeat_region 70560, .70692
/rpt_family="Alu"
repeat_region 72294, .72470
/rpt_family="Alu"
repeat_region complement(72800, .73095)
/rpt_family="Alu"
repeat_region 73424, .73695
/rpt_family="Alu"
repeat_region 73784, .74071
/rpt_family="Alu"
repeat_region 74652, .74935
/rpt_family="Alu"
repeat_region complement(75265, .75514)
/rpt_family="Alu"
repeat_region complement(76086, .76353)
/rpt_family="Alu"
repeat_region complement(78038, .78312)
/rpt_family="Alu"
repeat_region 78513, .78734
/rpt_family="Alu"
repeat_region 78866, .79180
/rpt_family="Alu"
repeat_region complement(79384, .79673)
/rpt_family="Alu"
repeat_region complement(81398, .81690)
/rpt_family="Alu"
repeat_region 82013, .82288
/rpt_family="Alu"
repeat_region complement(83419, .83600)
/rpt_family="Alu"
repeat_region complement(87320, .87624)
/rpt_family="Alu"
repeat_region complement(87710, .87987)
/rpt_family="Alu"
repeat_region complement(88186, .88472)
/rpt_family="Alu"
repeat_region 88822, .89049
/rpt_family="Alu"
repeat_region 89224, .89308
/rpt_family="Alu"
repeat_region complement(89729, .90487)
/rpt_family="Alu"
repeat_region 90504, .90618
/rpt_family="Alu"
repeat_region complement(93313, .93952)
/rpt_family="Alu"
repeat_region 96938, .97517
/rpt_family="Alu"
repeat_region complement(10419, .104514)
/rpt_family="Alu"
repeat_region complement(109738, .109979)
/rpt_family="Alu"
repeat_region 110022, .110141
/rpt_family="Alu"
repeat_region complement(111091, .112016)
```

```
Query Match 80.7%; Score 4034.2; DB 9; Length 196080;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 4486; Conservative 0; Mismatches 28; Indels 365; Gaps 7;

repeat_region /rpt_family="Alu"
113210, .113419
repeat_region /rpt_family="Al"
complement(114951, .115005)
/rpt_family="MIR"
115685, .115791
repeat_region /rpt_family="MIR"
116870, .117005
repeat_region /rpt_family="MIR"
complement(117877, .118172)
/rpt_family="Alu"
118352, .118600
repeat_region /rpt_family="Alu"
118679, .118813
repeat_region /rpt_family="Alu"
complement(121233, .121520)
/rpt_family="Alu"
122668, .123071

122 GCACCCACCATGGACACAGTATACCTATGTAAACCACTGCACCATCATATACCTAT 181
|||||
178042 GCACCCACCATGGACACAGTATACCTATGTAAACCACTGCACCATCATATACCTAT 178100
|||||
182 GTACCAACCTGTGTATCTGTCACACAGTATCCAGACTTTAGTGAAAAAAAGTG 241
|||||
178101 GTACCAACCT-GTACATTTCTGACACAGTATCCAGAAC-TTAAAGTAAAAAAAGTG 178158
|||||
242 GTGTGTAGAAAAATACCTGCAATCTGACATGTAAAGCTTATGATACATTTAGAGAGA 301
|||||
178159 GTGTGTAGAAAAATACCTGCAATCTGACATGTAAAGCTTATGATACATTTAGAGAGA 178218
|||||
302 GAGGTGACAGAAAGGAGGTGAGTGGGTTTAAACACAAAGTATTTAAAT 361
|||||
178219 GAGGTGACAGAAAGGAGGTGAGTGGGTTTAAACACAAAGTATTTAAAT 178278
|||||
362 CAGGCTCTGTGAACTTATGATCCCAAAACACATCTCATATCCCTGCAGTCTGA 421
|||||
178279 CAGGCTCTGTGAACTTATGATCCCAAAACACATCTCATATCCCTGCAGTCTGA 178338
|||||
422 TTAATAATGCAATCCCTTAAGGCCACAGACTCAACTGTGAGAAAGATCCAGAAACTG 481
|||||
178339 TTAATAATGCAATCCCTTAAGGCCACAGACTCAACTGTGAGAAAGATCCAGAAACTG 178398
|||||
482 CCCGTTTAAATTAACATTTGGGCGATCTTACGGCTCTAAAGACCAAGACCACTGCTGC 541
|||||
178399 CCCGTTTAAATTAACATTTGGGCGATCTTACGGCTCTAAAGACCAAGACCACTGCTGC 178458
|||||
542 CTAGAGCTGTGCTCTCTTCAATTAACAAATACAGAGAGTGTAGTAGACACCCACCA 601
|||||
178459 CTAGAGCTGTGCTCTCTTCAATTAACAAATACAGAGAGTGTAGTAGACACCCACCA 178518
|||||
602 CTTCACACAGCTTGAAGAGCCCTGAGTATGATGATGATTAATTAATTTATGATCA 661
|||||
178519 CTTCACACAGCTTGAAGAGCCCTGAGTATGATGATGATTAATTAATTTATGATCA 178578
|||||
662 CATGCTAGATTTTACACAGCTCCCGTGGGATCTGGGCAATTTATCCATATTCACAT 721
|||||
178579 CATGCTAGATTTTACACAGCTCCCGTGGGATCTGGGCAATTTATCCATATTCACAT 178638
|||||
722 GCGTGGCTGAGAGCGACAGCAATAAATCCAGAGGCTGTCTGTCAACCCCCACAGACTG 781
|||||
178639 GCGTGGCTGAGAGCGACAGCAATAAATCCAGAGGCTGTCTGTCAACCCCCACAGACTG 178698
|||||
782 ACCCCCTCCACAGCCCGGAGGCTTCTCTTCATCTCTCTGAAGCAATTAATGAT 841
|||||
178699 ACCCCCTCCACAGCCCGGAGGCTTCTCTTCATCTCTCTGAAGCAATTAATGAT 178758
|||||
842 GGGCCCTGCGAGCATCAAGCCAGAAATTAAGTATGATGATCAACGACGCCAATGACAG 901
|||||
```

Db 178759 GGGCCCTGCGCAGCCAAATCACAGCCAGAAATACGATGATGTGACACAGACCAATGAGG 178818
QY 902 CTCCTCGTCAGCATATGACGAAATCTGTCAATTTTACTAGAGGTGTGAAATCTCCCAAGCA 961
Db 178819 CTCCTCGTCAGCATATGACGAAATCTGTCAATTTTACTAGAGGTGTGAAATCTCCCAAGCA 961
QY 962 CACCATCTCTTTTTCAGATTAAGGACACAGGAGCTGAGAGAGAGAGCTGAACCTACCCGGGT 178878
Db 178879 CACCATCTCTTTTTCAGATTAAGGACACAGGAGCTGAGAGAGAGAGCTGAACCTACCCGGGT 1021
QY 1022 CACACACACAGAGTGGCGAAAGGCTGGGACCAAGAAACCCAGAGCTGTGATCTGCAGCCGGTA 178938
Db 178939 CACACACACAGAGTGGCGAAAGGCTGGGACCAAGAAACCCAGAGCTGTGATCTGCAGCCGGTA 1081
QY 1082 TTCATCTTTTCCATATAGCCACAGGGCTGTCAAAAGCCCCAGGCTTACTGACAGCCGGTA 178998
Db 178999 TTCATCTTTTCCATATAGCCACAGGGCTGTCAAAAGCCCCAGGCTTACTGACAGCCGGTA 1141
QY 1142 CCTCTCTGAGAGTTCCTGGCAGACAGAAATTAAGCTGACACACCCCTTAACCCCAAC 1201
Db 179059 CCTCTCTGAGAGTTCCTGGCAGACAGAAATTAAGCTGACACACCCCTTAACCCCAAC 1201
QY 1202 TCTCTCTGCAAGGCTTCAGGGGCTGAGAACTGCTGAGAGAGATCTTTAGCCCTGGAT 179118
Db 179119 TCTCTCTGCAAGGCTTCAGGGGCTGAGAACTGCTGAGAGAGATCTTTAGCCCTGGAT 1261
QY 1262 TTTAGGGCCATGATAGAGGGGGGTGTCCCTTAATTTCCAGCCCTGTCTCAGCCCAAC 179178
Db 179179 TTTAGGGCCATGATAGAGGGGGGTGTCCCTTAATTTCCAGCCCTGTCTCAGCCCAAC 1321
QY 1322 CCTCTCAAGAAATTAAGAGGGGCTATGAGGAGCTGTGAGGCTGTCTGAGCCCTTCTGACAG 179238
Db 179239 CCTCTCAAGAAATTAAGAGGGGCTATGAGGAGCTGTGAGGCTGTCTGAGCCCTTCTGACAG 1381
QY 1382 ATTACAAAGAGGACTTAACAAAGAGCTCTTTGTGAGAGCTGTGCTGATAGGAGTCAAG 179298
Db 179299 ATTACAAAGAGGACTTAACAAAGAGCTCTTTGTGAGAGCTGTGCTGATAGGAGTCAAG 1441
QY 1442 TGACGGCGGCTAGACACTACGCTGGGAGGTGCCAGCTCTTAAGAGTGGGACAGGGCACTG 179358
Db 179359 TGACGGCGGCTAGACACTACGCTGGGAGGTGCCAGCTCTTAAGAGTGGGACAGGGCACTG 1501
QY 1502 GCCACAGAGTCCAGAGGAGTCCACACAGCTTAAGTGCACAGCTTCTGTGGGATCATGG 179418
Db 179419 GCCACAGAGTCCAGAGGAGTCCACACAGCTTAAGTGCACAGCTTCTGTGGGATCATGG 1561
QY 1562 ACCACACGGAACCCACCTGTGAGTACAAAGTCCACAGTGGAGTGGAGCTGGGCTGGCTTT 179478
Db 179479 ACCACACGGAACCCACCTGTGAGTACAAAGTGGAGTGGAGTGGAGCTGGGCTGGCTTT 1621
QY 1622 GAGGCTTTACAGGGTTGATGGCCATCTGTGCGTATTTGTGTGGATATGACACACAGGCA 179538
Db 179539 GAGGCTTTACAGGGTTGATGGCCATCTGTGCGTATTTGTGTGGATATGACACACAGGCA 1681
QY 1682 GCACATGCGCAGAGTGTGTGGCAGCTGTGTGCTGTGCAAAATGCCCCTGAGGTGGGAATGA 179598
Db 179599 GCACATGCGCAGAGTGTGTGGCAGCTGTGTGCTGTGCAAAATGCCCCTGAGGTGGGAATGA 1741
QY 1742 GCTTGTGTGCATCAGGACGACAGCCAGCTGTGTGCTGTGAGAAATGCCCCTGAGGTGGGAATGA 179658
Db 179659 GCTTGTGTGCATCAGGACGACAGCCAGCTGTGTGCTGTGAGAAATGCCCCTGAGGTGGGAATGA 1801
QY 1802 ACATATGAGAGGGGATCATATCATCATCAAAAATATTATTATAGAGCTCCCTTAAAAAAAAG 179715
Db 179716 ACATATGAGAGGGGATCATATCATCATCAAAAATATTATTATAGAGCTCCCTTAAAAAAAAG 1861
QY 1862 AAGGTCTCTTTTCGATTAAGAAAGGAGAGAGGGGTTGTCTTATTAATTAAGG 179738
Db 179739 AAGGTCTCTTTTCGATTAAGAAAGGAGAGAGGGGTTGTCTTATTAATTAAGG 1921
QY 1922 AGGAGCGGCCCTCAAAAAATTAAGGAGGAGAGAGACCCAGACCCCGTGGTGTGTGTT 179738
Db 179739 AGGAGCGGCCCTCAAAAAATTAAGGAGGAGAGAGAGACCCAGACCCCGTGGTGTGTGTT 1981
179738 179738

QY	1982	TCACGGGGAGCTCGAACCTTTAGAGGAGCGTGGGAACCCGCTGTAATTCAGGCTCT	2041
Db	179739	-----	179739
QY	2042	CGAGAGAAAGGAGCGCGCCGCCAAAAAATATCCCTCCCGGGCATTAAGAAATGTGGCC	2101
Db	179739	-----	179739
QY	2102	TCCTCAAAAAGATGAAGAGGAAGCCGGAGTGTATGTGTAATTTTAAAACTCCAG	2161
Db	179739	-----	179739
QY	2162	GTAGNNNNNNNNNNNGCTTCAGTAATTTTATTTGAGCGCTTCTACGAGAACACAGAAG	2221
Db	179739	TCCTTCAGTAATTTTATTTGAGCGCTTCTACGAGAACACAGAAG	179784
QY	2222	GAGCTTCATTTCTGAGAGGAAACAGCGAGAAACAGCGAGATTCCTGTATTAATTCAA	2281
Db	179785	GAGCTTCATTTCTGAGAGGAAACAGCGAGAAACAGCGAGATTCCTGTATTAATTCAA	179844
QY	2282	GTAAGTAAATGTGCTCTCAGAAATATCAAGCAAGGTGAGAGACACAGACAGCGGTGG	2341
Db	179845	GTAAGTAAATGTGCTCTCAGAAATATCAAGCAAGGTGAGAGACACAGACAGCGGTGG	179904
QY	2342	CAGTGGGCTGTATTTCCAGGTGGATGGTGGGAACTCCTTTCTAAAGGAACTTGA	2401
Db	179905	CAGTGGGCTGTATTTCCAGGTGGATGGTGGGAACTCCTTTCTAAAGGAACTTGA	179964
QY	2402	GTGGGAAGGAACATCGACAGATATCTCAGGAAGACCTTCTCCAGCGAGGAAGATCAGAG	2461
Db	179965	GTGGGAAGGAACATCGACAGATATCTCAGGAAGACCTTCTCCAGCGAGGAAGATCAGAG	180024
QY	2462	GTGGAAAGCCCTTGAGACCACATTCAGTAATCATCTTGAGCATCTCTACCAAGCTAG	2521
Db	180025	GTGGAAAGCCCTTGAGACCACATTCAGTAATCATCTTGAGCATCTCTACCAAGCTAG	180084
QY	2522	TTCCATTAATGGAATGGGAATATGCGGTGGAGACAGGCGTGCCTGCTCCATCACTTC	2581
Db	180085	TTCCATTAATGGAATGGGAATATGCGGTGGAGACAGGCGTGCCTGCTCCATCACTTC	180144
QY	2582	TCACACTAGGCTGTGGAGAGACTTGGAGCTTAACGAACAGATGGGTGGAACACTG	2641
Db	180145	TCACACTAGGCTGTGGAGAGACTTGGAGCTTAACGAACAGATGGGTGGAACACTG	180204
QY	2642	CCTAGCCAGAGAGACTTGAGCTTAATGTGTAGACATTCCTGCTGTACTGCTTTGTAT	2701
Db	180205	CCTAGCCAGAGAGACTTGAGCTTAATGTGTAGACATTCCTGCTGTACTGCTTTGTAT	180264
QY	2702	TGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	2761
Db	180265	TGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	180324
QY	2762	CTTGAAGTCAATGGCGATCTCAGCTCAGCTCAGCACTCCACCTCTGGGATCAAGGAT	2821
Db	180325	CTTGAAGTCAATGGCGATCTCAGCTCAGCTCAGCACTCCACCTCTGGGATCAAGGAT	180384
QY	2822	TCACCTGCTCAGCTCTCTGAGTAGTGGGATTAACAGGCAACCGGACACAGCTGTAA	2881
Db	180385	TCACCTGCTCAGCTCTCTGAGTAGTGGGATTAACAGGCAACCGGACACAGCTGTAA	180444
QY	2882	TTTTTTTGTATTTTGTAGAGACAGGTTTCAACATGTTGGCCAGGCTGTGCAACT	2941
Db	180445	TTTTTTTGTATTTTGTAGAGACAGGTTTCAACATGTTGGCCAGGCTGTGCAACT	180504
QY	2942	CCGAGACCTTAAGTATCCACCTGCTGCACTCCCAAGTGGTGGGATTAATAGCATAG	3001
Db	180505	CCGAGACCTTAAGTATCCACCTGCTGCACTCCCAAGTGGTGGGATTAATAGCATAG	180564
QY	3002	CCACGTGCCCCAGTATTAAGAAAGTTAAAGGCATGGCAATGGCAATGCAACGCTATCTAC	3061
Db	180565	CCACGTGCCCCAGTATTAAGAAAGTTAAAGGCATGGCAATGGCAATGCAACGCTATCTAC	180624

Db	181697	GCTAATTTTTTTTTTTTTTTCTGATTTTTTACTAGACAGAGGTTTCATCATGTTGGCCAGG	181756
Qy	4202	CTGGTCCTTGAACCCCGACCTCAAGATGATCCACCCACTCGGCTCCCAAAAGTCTGGGA	4281
Db	181757	CTGGTCCTTGAACCCCGACCTCAAGATGATCCACCCACTCGGCTCCCAAAAGTCTGGGA	181816
Qy	4262	TTACAGGTGTCAAGCCACATATGCACAGCCCATGATGATTTTTTAAATATATTTTTTAA	4321
Db	181817	TTACAGGTGTCAAGCCACATATGCACAGCCCATGATGATTTTTTAAATATATTTTTTAA	181876
Qy	4322	TTTAAATGTTTATCTTAAGCCAGTACAGTACTCGCGTCTGTATTCACAGCACTTTGAG	4381
Db	181877	TTTAAATGTTTATCTTAAGCCAGTACAGTACTCGCGTCTGTATTCACAGCACTTTGAG	181936
Qy	4382	GGGCCAAGGTGCGGGGATCACTTGAGCCTTGAGGATTCAGCGTGGGGCAATATAGAGACC	4441
Db	181937	GGGCCAAGGTGCGGGGATCACTTGAGCCTTGAGGATTCAGCGTGGGGCAATATAGAGACC	181996
Qy	4442	CCGTCCTCAACAAAAATTTAAAAAATTTAGCTGGAGTGTGSCATTTTCCCTGTGTCCCA	4501
Db	181997	CCGTCCTCAACAAAAATTTAAAAAATTTAGCTGGAGTGTGSCATTTTCCCTGTGTCCCA	182056
Qy	4502	GCTACTTGGGAAGCTAGGCTGGGATGGCTTAAGCCTTGAGGTCAGAGGCTGCAGCTGA	4561
Db	182057	GCTACTTGGGAAGCTAGGCTGGGATGGCTTAAGCCTTGAGGTCAGAGGCTGCAGCTGA	182116
Qy	4562	GCTATGATACACACCACTGCACTTCAGCTGAGTGTAGAGGCTATCTCAAAAGCAACAAA	4621
Db	182117	GCTATGATACACACCACTGCACTTCAGCTGAGTGTAGAGGCTATCTCAAAAGCAACAAA	182176
Qy	4622	TATGTTTATCTAAAGTAAAGGTAAATACAGAAATATATGATATGATTTTAAATTGAA	4681
Db	182177	TATGTTTATCTAAAGTAAAGGTAAATACAGAAATATATGATATGATTTTAAATTGAA	182236
Qy	4682	AAAGCATTAATGATTAACATGATTTGTAATAATATCAATACATGAAATTTCTTGTTCTTA	4741
Db	182237	AAAGCATTAATGATTAACATGATTTGTAATAATATCAATACATGAAATTTCTTGTTCTTA	182296
Qy	4742	ATAAAGCTAGCAACAAGGACATTTGGTTTAACTAGGGCACCAAGTACTTTAAAAAAA	4801
Db	182297	ATAAAGCTAGCAACAAGGACATTTGGTTTAACTAGGGCACCAAGTACTTTAAAAAAA	182356
Qy	4802	GTTAGGGCCAGCCACAGGGGCTCACACCTGTATATCCACACTTTGGGAGCCCAAGCAG	4861
Db	182357	GTTAGGGCCAGCCACAGGGGCTCACACCTGTATATCCACACTTTGGGAGCCCAAGCAG	182416
Qy	4862	GAGGATCACTTGAAGCCAGAGATTTTGAACCTGAGACACATATAGGAGATCTGATCTTGT	4921
Db	182417	GAGGATCACTTGAAGCCAGAGATTTTGAACCTGAGACACATATAGGAGATCTGATCTTGT	182476
Qy	4922	CTCTATTAATAAATTTAAAAAATTTGGCTAGGACCTTTGGCTTACACCCGTATCCAGCACT	4981
Db	182477	CTCTATTAATAAATTTAAAAAATTTGGCTAGGACCTTTGGCTTACACCCGTATCCAGCACT	182536
Qy	4982	TTGGGAGGCCGAGCGCGGT 5000	
Db	182537	TTGGGAGGCCGAGCGCGGT 182555	
RESULT 4			
AC087451/c			
LOCUS			
DEFINITION Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT			
ACCESSION AC087451			
VERSION AC087451.2 GI:14595836			
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE Homo sapiens.			
ORGANISM Homo sapiens.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE 1 (bases 1 to 163915)			

[illegible]

QY	4818	GGGGCTCACACCTGTATATCCAGACTTTGGAGAGCCAGGACGAGATCATCTTGAGCC	4877
Db	160377	GGGGCTCACACCTGTATATCCAGACTTTGGAGAGCCAGGACGAGATCATCTTGAGCC	160318
QY	4878	CAGAGCTTTGAGACCTGAGCAATATAGGAGATCTGTATCTCTATATAAAATTA	4937
Db	160317	CAGAGCTTTGAGACCTGAGCAATATAGGAGATCTGTATCTCTATATAAAATTA	160328
QY	4938	AAATTTGGCTAGGCCCTTTGGCTTAACCCGTAATCCAGCACTTTGGAGAGCCGAGCC	4997
Db	160257	AAATTTGGCTAGGCCCTTTGGCTTAACCCGTAATCCAGCACTTTGGAGAGCCGAGCC	160198
QY	4998	GCT 5000	
Db	160197	GCT 160195	
RESULT 5	AP003733		
LOCUS	166867	bp	linear PRI 08-DEC-2001
DEFINITION	Homo sapiens genomic DNA, chromosome 11q, clone:RP11-810P12,		
ACCESSION	AP003733		
VERSION	AP003733.4		
KEYWORDS	GI:17426128		
SOURCE	HTG.		
ORGANISM	Homo sapiens DNA, clone:RP11-810P12.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.		
JOURNAL	Published Only in Database (2001)		
REFERENCE	2 (bases 1 to 166867)		
AUTHORS	Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Shutoh-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan		
COMMENT	(E-mail: hattori@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)		
FEATURES	On Dec 7, 2001 this sequence version replaced gi:17026124.		
SOURCE	Location/Qualifiers		
	1. 166867		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q"		
	/clone="RP11-810P12"		
BASE COUNT	44705 a 38465 c 39268 g 44429 t		
ORIGIN			
Query Match	65.18; Score 3254.6; DB 9; Length 166867;		
Best Local Similarity	90.58; Pred. No. 0;		
Matches 3694; Conservative 0; Mismatches 24; Indels 365; Gaps 6;			
QY	921	GAATTCGTCAATTTACTAGGGTGATGAATATCCCAAGCAACGATCCTTTTCAGATTA	980
Db	1	GAATTCGTCAATTTACTAGGGTGATGAATATCCCAAGCAACGATCCTTTTCAGATTA	980
QY	981	GGGCACTGAGGGTGAGAGAGAGCTGAAAACCTAACCCGGGGTACCAACAGAGTGGCAA	1040
Db	61	GGGCACTGAGGGTGAGAGAGAGCTGAAAACCTAACCCGGGGTACCAACAGAGTGGCAA	1040
QY	1041	GGCTGGAGCAACCAAGCAAGCTGTGACTGACGCCGGATTCATTTTCCATAGCC	1100
Db	121	GGCTGGAGCAACCAAGCAAGCTGTGACTGACGCCGGATTCATTTTCCATAGCC	1100
QY	1101	ACAGGGCTGTAAAGACCCCAAGGGCTAGTGCAGAGGCTCCCTTCCGGAGAGTTCCTG	1160

Db 181 ACAGGGCTGTCAAAAGACCAGGGCCCTAGTCAGAGGCTCCTCTTCTGAGAGTTCCTG 240
QY 1161 GCACAGAGTTGAAGCTCAGACACAGCCCTTAACCCCAACTCTCTCTGCAAGGCTCAG 1220
Db 241 GCACAGAGTTGAAGCTCAGACACAGCCCTTAACCCCAACTCTCTCTGCAAGGCTCAG 300
QY 1221 GGGTCAGAACTAGTGTGAGACAGATCCTTAAAGCTCTGATTTTAAAGGCGCATGTAGAG 1280
Db 301 GGGTCAGAACTAGTGTGAGACAGATCCTTAAAGCTCTGATTTTAAAGGCGCATGTAGAG 360
QY 1281 GGGTGTGGCCCTAAATTCAGCCCTGTGTCAAGCCCAACACCTCCCAAGAAATTAAG 1340
Db 361 GGGTGTGGCCCTAAATTCAGCCCTGTGTCAAGCCCAACACCTCCCAAGAAATTAAG 420
QY 1341 GGGGCGATGGCGGCTGTGCTAGCCCTTCTCTGAGCAGATTAAGAGAGGACTAAG 1400
Db 421 GGGGCGATGGCGGCTGTGCTAGCCCTTCTCTGAGCAGATTAAGAGAGGACTAAG 480
QY 1401 ACAAGAGCTCCTTTGTGAGAGTCTGTGAGGATCAAGTACAGGCGGCTCAGACATC 1460
Db 481 ACAAGAGCTCCTTTGTGAGAGTCTGTGAGGATCAAGTACAGGCGGCTCAGACATC 540
QY 1461 AGCTGGGCGAGTGGCAGCCCTTAAGAGTGGGCGAGGCGACTGGCCACAGAGTCCAGGAG 1520
Db 541 AGCTGGGCGAGTGGCAGCCCTTAAGAGTGGGCGAGGCGACTGGCCACAGAGTCCAGGAG 600
QY 1521 TCCACACAGCTAGTGGCGAGACCTTCTGTGGGATCATGGAGCCCACTCCGGAACCCAC 1580
Db 601 TCCACACAGCTAGTGGCGAGACCTTCTGTGGGATCATGGAGCCCACTCCGGAACCCAC 660
QY 1581 TGTGAGTACAGAGTGGCGAGTGGGCTGTGGGCTTTAGAGCCTTCAGGCTTGAT 1640
Db 661 TGTGAGTACAGAGTGGCGAGTGGGCTGTGGGCTTTAGAGCCTTCAGGCTTGAT 720
QY 1641 GGCATCTTGGCTATTTGTGGGATGATGACACACAGCAGCAGATGGGCGAGTGTG 1700
Db 721 GGCATCTTGGCTATTTGTGGGATGATGACACACAGCAGCAGATGGGCGAGTGTG 780
QY 1701 GGCACCTGTGTGTGTGCAAAATGCCCTGAGGTGGAAATGAGCTTGTGTGATCAGAG 1760
Db 781 GGCACCTGTGTGTGTGCAAAATGCCCTGAGGTGGAAATGAGCTTGTGTGATCAGAG -- 838
QY 1761 GCAGAGCAGCAGTGTGGCTGACAGAAACACACAGGAAAGATGGAGGGGCTATCA 1820
Db 839 -CAGAGCCAGCAGTGTGGCTGACAGAAACACACAGGAAAGATGGAGGGGCTATCA 897
QY 1821 TCACAGCAAAATTTATAGAGCTCCCTTAAAAAAGAGAGTCTCTTTCGATTA 1880
Db 898 TCAC----- 901
QY 1881 GAAAGAAAGGAGAGAGGGGTTTCTTATAATAATAAGGAGAGCCGCCCTCAAAA 1940
Db 902 ----- 901
QY 1941 ATAAAGGAGGAGAGCAGCAGACCCCGTGGGTTGTGTTCAGGGGAGGCTCGAAC 2000
Db 902 ----- 901
QY 2001 CTTTAGAGGAGCCTGGGAGAACCGCTGATTCAGGCCCTCGAGAGAAAGAGAGCGCC 2060
Db 902 ----- 901
QY 2061 GCCCAAAAATATCCCTCCCGGGGATTAAGAAATGGTGGCTCTCTCAAAAAGATGAAGA 2120
Db 902 ----- 901
QY 2121 GGAAGCGAGTGTATGTGTGATATTTTAAACTCAGGTAGNNNNNNNNNTGCTT 2180
Db 902 -----TGCCT 906
QY 2181 CAGTAAATTTTATGAGCGCTTCTAGAGACAGAGAGAGACTTCCATTCAGAGAG 2240
Db 907 CAGTAAATTTTATGAGCGCTTCTAGAGAGACAGAGAGAGACTTCCATTCAGAGAG 966

QY 2241 GAAACAGCAGAGAAACAGCAGATATCTGTATTAATTTCAAGTATGATAGGCTCT 2300
Db 967 GAAACAGCAGAGAAACAGCAGATATCTGTATTAATTTCAAGTATGATAGGCTCT 1026
QY 2301 AGAAATATCAGCAAGGTGAGAGACACAGAGCAGCCGGTGGAGTGGGCTATTTCCA 2360
Db 1027 AGAAATATCAGCAAGGTGAGAGACACAGAGCAGCCGGTGGAGTGGGCTATTTCCA 1086
QY 2361 GGTGGATGGTGGGAACTCTTTCTAAAGGAAACCTGGAGTGGGAAAGCAATGAG 2420
Db 1087 GGTGGATGGTGGGAACTCTTTCTAAAGGAAACCTGGAGTGGGAAAGCAATGAG 1146
QY 2421 GTATCTCAGAGAGACTTCTTCAGCAGAGAAAGATCAGAGGTGAGAAAGCCCTGAGCC 2480
Db 1147 GTATCTCAGAGAGACTTCTTCAGCAGAGAAAGATCAGAGGTGAGAAAGCCCTGAGCC 1206
QY 2481 ACCATTCAAGTAAACATTTGAGATCTCTACAGCTAGTTCATTTATGGAAATGGA 2540
Db 1207 ACCATTCAAGTAAACATTTGAGATCTCTACAGCTAGTTCATTTATGGAAATGGA 1266
QY 2541 ATATGGTGGAGACAGGCTGCTGCTCCCTTCATATCTCAGACTAGGCTGCTGAG 2600
Db 1267 ATATGGTGGAGACAGGCTGCTGCTCCCTTCATATCTCAGACTAGGCTGCTGAG 1326
QY 2601 AGAGCTTGGAGCTTACGAAACAGATGGGCTGAGAACACTGCTAGCCAGAGACTGA 2660
Db 1327 AGAGCTTGGAGCTTACGAAACAGATGGGCTGAGAACACTGCTAGCCAGAGACTGA 1386
QY 2661 GCTTATGTATGAGATCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2720
Db 1387 GCTTATGTATGAGATCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1446
QY 2721 TTTATGATCTTAAAGACAGATTTTGTCTCTTACCAGGCTTGAAGTGAAGGCTGA 2780
Db 1447 TTTATGATCTTAAAGACAGATTTTGTCTCTTACCAGGCTTGAAGTGAAGGCTGA 1506
QY 2781 TCTCAGCTCAGTGCAGCTCAGCTCCCGGATCAAGGATTCCTGCGCCAGCCTTCT 2840
Db 1507 TCTCAGCTCAGTGCAGCTCAGCTCCCGGATCAAGGATTCCTGCGCCAGCCTTCT 1566
QY 2841 GAGTACGTGGATTAACAGGACCCGACACAGCCTGTGATTAATTTTGTATTTTATGA 2900
Db 1567 GAGTACGTGGATTAACAGGACCCGACACAGCCTGTGATTAATTTTGTATTTTATGA 1626
QY 2901 GAGACAGGCTTACCAATGTGGCAGAGCTGTCCCACTCCGACTTATAGTATGATCA 2960
Db 1627 GAGACAGGCTTACCAATGTGGCAGAGCTGTCCCACTCCGACTTATAGTATGATCA 1686
QY 2961 CCTGCTCGACTTCCCAAGTGTGGGATTAAGCATGAGCAGCAGTCCGACAGTGAATTA 3020
Db 1687 CCTGCTCGACTTCCCAAGTGTGGGATTAAGCATGAGCAGCAGTCCGACAGTGAATTA 1746
QY 3021 TAGAAATTAAGGCAATGAGCAATGACAGCAGCTATCTAGCTTCCCTGCAAGCA 3080
Db 1747 TAGAAATTAAGGCAATGAGCAATGACAGCAGCTATCTAGCTTCCCTGCAAGCA 1806
QY 3081 AGGCGACCTCTGGGCTCAGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3140
Db 1807 AGGCGACCTCTGGGCTCAGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1866
QY 3141 AGGCGCTTGGAGACCACTTATCCACTCTAGAGTCCCTATGAGAGTGTGAGTTCAG 3200
Db 1867 AGGCGCTTGGAGACCACTTATCCACTCTAGAGTCCCTATGAGAGTGTGAGTTCAG 1926
QY 3201 AGCAGGAAAGGCTCTCAGAGGCTCTGACAGGCGCTCTGATCCCTTCAAAACCCCAATC 3260
Db 1927 AGCAGGAAAGGCTCTCAGAGGCTCTGACAGGCGCTCTGATCCCTTCAAAACCCCAATC 1986
QY 3261 GGTGTCCCTCTTACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3320
Db 1987 GGTGTCCCTCTTACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2046

QY 3321 ATCACTTACACAGCAAGTGGCTAATGCCCCGCTTAGGCTCTCTCCGCTCTGCTG 3380
 Db 2047 ATCACTTACACAGCAAGTGGCTAATGCCCCGCTTAGGCTCTCTCCGCTCTGCTG 2106
 QY 3381 TGGCGGGGGGCGACATCTACAGCTCTATATGCGAGTTCCTAATCTCTCTGCTG 3440
 Db 2107 TGGCGGGGGGCGACATCTACAGCTCTATATGCGAGTTCCTAATCTCTCTGCTG 2166
 QY 3441 TACTACATCATCGCTTATTTATAGTAAAGCTGCGAGGCTGGGCGGGGGGCTGGG 3500
 Db 2167 TACTACATCATCGCTTATTTATAGTAAAGCTGCGAGGCTGGGCGGGGGGCTGGG 2226
 QY 3501 AAGGATGTGGCTGGGCTGGGAGCTCTGGGAGCTCTGGGAGCTCTGGGAGCTCAAGGC 3560
 Db 2227 AAGGATGTGGCTGGG-----GCTGGAGCTCTCTGGGAGCTCTGGGAGCTCAAGGC 2279
 QY 3561 CCAGTGCACACAGTCCCTCAACACTAGCTGGGCTCTGACAGCTCTGGGAGCTCAAGGC 3620
 Db 2280 CCAGTGCACACAGTCCCTCAACACTAGCTGGGCTCTGACAGCTCTGGGAGCTCAAGGC 2339
 QY 3621 GCTGAGGCTGCGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAG 3680
 Db 2340 GCTGAGGCTG-----GGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGG 2398
 QY 3681 GCGAGGCGCAGAGGCTTTTACGACACCTCTCCCAACCCAGGAGAGCTGGGAGCTGGGAG 3740
 Db 2399 GCGAGGCGCAGAGGCTTTTACGACACCTCTCCCAACCCAGGAGAGCTGGGAGCTGGGAG 2458
 QY 3741 GCGCTTGTCTGCGCCACTCTACAGCTCTGCTTTTACTGAATCCACACAGCTCAAGGC 3800
 Db 2459 GCGCTTGTCTGCGCCACTCTACAGCTCTGCTTTTACTGAATCCACACAGCTCAAGGC 2518
 QY 3801 CCACATGTATATTAAAA--AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 3858
 Db 2519 CCACATGTATATTAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 2578
 QY 3859 CTGTGTGTCCAGGCTGCTCCAGCTCTAGGCTCAAGCAATCCCTCTCTGAGCTG 3918
 Db 2579 CTGTGTGTCCAGGCTGCTCCAGCTCTAGGCTCAAGCAATCCCTCTCTGAGCTG 2638
 QY 3919 CCAGAGGCGCTGGATTACAGAGTGTGAGCTCTGACCTTGACCAACACATGATG 3977
 Db 2639 CCAGAGGCGCTGGATTACAGAGTGTGAGCTCTGACCTTGACCAACACATGATG 2698
 QY 3978 TTTTGT 4037
 Db 2699 TTTTGT 2758
 QY 4038 GGGGCAATCTGGCTCAGCTGTAACTCTGCTCCAGGCTGCAAGGATCTCTGCTG 4097
 Db 2759 GGGGCAATCTGGCTCAGCTGTAACTCTGCTCCAGGCTGCAAGGATCTCTGCTG 2818
 QY 4098 GCGCTCTGAGTGTGAATATAGGACACACACACACACACACACACACACACACAC 4157
 Db 2819 GCGCTCTGAGTGTGAATATAGGACACACACACACACACACACACACACACACAC 2878
 QY 4158 TTTGTATATTATAGAGAGAGAGGTTTCAATCATGTTGGCCAGGCTGTGTAACCC 4217
 Db 2879 TTTGTATATTATAGAGAGAGGTTTCAATCATGTTGGCCAGGCTGTGTAACCC 2938
 QY 4218 GACCTCAAGTGTACCCACACCTGGGCTCCCAAGGCTGGGATTAACAGTGTACACCA 4277
 Db 2939 GACCTCAAGTGTACCCACACCTGGGCTCCCAAGGCTGGGATTAACAGTGTACACCA 2998
 QY 4278 CCATGACACACACATGTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4337
 Db 2999 CCATGACACACACATGTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 3058
 QY 4338 AGGCAAGTGTACAGTGTGCTGTATATCCAGACACTTTGAGGGGCAAGTGTGGG 4397
 Db 3059 AGGCAAGTGTACAGTGTGCTGTATATCCAGACACTTTGAGGGGCAAGTGTGGG 3118
 QY 4398 ATCACTTACAGCTGGGAGCTGCGGCAATATGTAGACCCCGCTCTACCAAAAA 4457

Db 3119 ATCACTTACAGCTGGGAGCTTACAGCTGGGAGCAATGTAGAGACCCCTCTCTACAAAA 3178
 QY 4458 TTTAAAAAATTTACTGGGAGTGTGGCATTTGCTGTGCTCCAGCTACTTGGAGCTG 4517
 Db 3179 TTTAAAAAATTTACTGGGAGTGTGGCATTTGCTGTGCTCCAGCTACTTGGAGCTG 3238
 QY 4518 AGGTGTGGGAGTGTGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 4577
 Db 3239 AGGTGTGGGAGTGTGGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 3298
 QY 4578 TGCATCTTACGCTGAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 4637
 Db 3299 TGCATCTTACGCTGAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 3358
 QY 4638 GGTAAAT 4697
 Db 3359 AATATAAT 3418
 QY 4698 CATGATTTGTAAAT 4757
 Db 3419 CATGATTTGTAAAT 3478
 QY 4758 GGCACATTTGCTTTTACTAGGAGCAGCAAGGCTACTTTAAAAAAGTTAGGCGCAGCA 4817
 Db 3479 GGCACATTTGCTTTTACTAGGAGCAGCAAGGCTACTTTAAAAAAGTTAGGCGCAGCA 3538
 QY 4818 GGGGCTCACACCTGTAT 4877
 Db 3539 GGGGCTCACACCTGTAT 3598
 QY 4878 CAGAGTGTAGAGCTGTGAGCAATAGGAGATCTGTATCTGTATATATATATATAT 4937
 Db 3599 CAGAGTGTAGAGCTGTGAGCAATAGGAGATCTGTATCTGTATATATATATATAT 3658
 QY 4938 AAAATGGCTAGGCTGCTTTGGCTTACACCTGTATATATATATATATATATATAT 4997
 Db 3659 AAAATGGCTAGGCTGCTTTGGCTTACACCTGTATATATATATATATATATATAT 3718
 QY 4998 GGT 5000
 Db 3719 GGT 3721

RESULT 6
 AC084857
 LOCUS
 DEFINITION Homo sapiens chromosome 11, clone CTD-3231N5 map 11, WORKING DRAFT
 AC084857
 ACCESSION
 VERSION AC084857.1 GI:11276215
 KEYWORDS HTG, HTGS, PHASE1, HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 Birten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Barn, N., Bostlen, V., Bida, F., Boguslavsky, L.,
 Bouhgalter, B., Brown, A., Burkett, G., Campoplano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collumore, A., Cooke, P.,
 Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
 Iliav, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, M.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,
 Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrum, J., Meneus, L., Mllove, T., Mlenga, Y.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,

TITLE
JOURNAL
COMMENT

O'Donnell, P., O'Neill, D., Oliveira, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Reback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schuer, S., Severy, P.,
Sougen, C., Spencer, B., Stange, Thomas, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
Tirell, A., Travers, M., Trillio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wymann, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: 3231_N.5
Center clone name: 3231_N.5

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 126761 bases at least Q40
Consensus quality: 129952 bases at least Q30
Consensus quality: 131351 bases at least Q20
Insert size: 128000; agarose-IP
Insert size: 132483; sum-of-ctrls
Quality coverage: 6.1 in Q20 bases; sum-of-ctrls
Quality coverage: 5.9 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1459: contig of 1459 bp in length
* 1460 1559: gap of 100 bp
* 1560 2520: contig of 961 bp in length
* 2521 2620: gap of 100 bp
* 2621 6334: contig of 3714 bp in length
* 6335 6434: gap of 100 bp
* 6435 11913: contig of 5479 bp in length
* 11914 12013: gap of 100 bp
* 12014 20756: contig of 8743 bp in length
* 20757 20856: gap of 100 bp
* 20857 29735: contig of 8879 bp in length
* 29736 29835: gap of 100 bp
* 29836 38645: contig of 8810 bp in length
* 38646 38745: gap of 100 bp
* 38746 52193: contig of 13448 bp in length
* 52194 52293: gap of 100 bp
* 52294 65209: contig of 12916 bp in length
* 65210 65309: gap of 100 bp
* 65310 79796: contig of 14487 bp in length
* 79797 79896: gap of 100 bp
* 79897 95896: contig of 16000 bp in length
* 95897 95996: gap of 100 bp
* 95997 113419: contig of 17423 bp in length
* 113420 113519: gap of 100 bp
* 113520 133683: contig of 20164 bp in length.

FEATURES
SOURCE
1. 133683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"

/clone="CTD-3231N5"
/clone.lib="CTD Human BAC"
1. 1459
/note="assembly-fragment"
1560. 2520
/note="assembly-fragment"
2621. 6334
/note="assembly-fragment"
6435. 11913
/note="assembly-fragment"
12014. 20756
/note="assembly-fragment"
20857. 29735
/note="assembly-fragment"
29836. 38645
/note="assembly-fragment"
38746. 52193
/note="assembly-fragment"
52294. 65209
/note="assembly-fragment"
65310. 79796
/note="assembly-fragment"
79897. 95896
/note="assembly-fragment"
95997. 113419
/note="assembly-fragment"
113520. 133683
/note="assembly-fragment"
BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
ORIGIN

Query Match 62.8%; Score 3142; DB 2; Length 133683;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3609; Conservative 0; Mismatches 140; Indels 373; Gaps 7;
QY 879 ATGTACACGACGACCAATCAGAGCTCTGTCAGATATGCAATTCGTTCATTTTACT 938
DB 79796 ANNN 79855
QY 939 AGGCGATGAATTCCTCCAGCAACACATCTTTTCAGATAGGCGACTGAGCTGAG 998
DB 79856 NNN 79915
QY 999 AGGAGCTGAACCTACCGGGGTACACACAGCGTGGCAAGCTGGACCAAGAACCA 1058
DB 79916 AGGAGCTGAACCTACCGGGGTACACACAGCGTGGCAAGCTGGACCAAGAACCA 79975
QY 1059 GAGCTGTGACTGACGCGGCTATTCATTTTCCATAGCCACAGGCGTCAAGAGC 1118
DB 79976 GAGCTGTGACTGACGCGGCTATTCATTTTCCATAGCCACAGGCGTCAAGAGC 80035
QY 1119 CCAGGCGCTATTCAGAGGCTCTCTCTCTGAGAGTTCTTGGCACAGAGTTGAAGTC 1178
DB 80036 CCAGGCGCTATTCAGAGGCTCTCTCTCTGAGAGTTCTTGGCACAGAGTTGAAGTC 80095
QY 1179 AGCAGAGCCCTTACCCCACTCTCTCTCAAGGCGTGGGCTGGAACACTGGGG 1238
DB 80096 AGCAGAGCCCTTACCCCACTCTCTCTCAAGGCGTGGGCTGGAACACTGGGG 80155
QY 1239 AGCAGATCTTTAGCTCTGATTTTAGGGCCATGTAAGAGGGGTGTGGCTTAATTC 1298
DB 80156 AGCAGATCTTTAGCTCTGATTTTAGGGCCATGTAAGAGGGGTGTGGCTTAATTC 80215
QY 1299 CAGCGCTGGTCTGAGCCCAACCTCTCAAGAAATTTAGAGGGCGCATGGCCAGCTG 1358
DB 80216 CAGCGCTGGTCTGAGCCCAACCTCTCAAGAAATTTAGAGGGCGCATGGCCAGCTG 80275
QY 1359 TGCTAGCGTGTCTTCTGAGAGATTAAGAGAGGATTAAGAGAGGACTCTTGTGG 1418
DB 80276 TGCTAGCGTGTCTTCTGAGAGATTAAGAGAGGATTAAGAGAGGACTCTTGTGG 80335
QY 1419 AGCTCTGGCTTAGGAGATCAATGACGGGCGTCACTACCTGCGGAGTCCAGCC 1478
|||||

Db 80336 AGGTCTGCGTTAGGGAGTCAAGTGAAGCGGCGGCTCAGACATCCAGCTGGGCACTGCCAGCC 80395

QY 1479 TCTAAGAGTGGGAGGGGCACTGGGCCACAGAGTCCAGAGAGTCCACAGCTATGTCG 1538

Db 80396 TCTAAGAGTGGGAGGGGCACTGGGCCACAGAGTCCAGAGGAGTCCACAGCTATGTCG 80455

QY 1539 CAGACCTTCTGTGGATCATCGGACCACCTGMAACCCACCTGTGAGTACAAAGTGCC 1598

Db 80456 CAGACCTTCTGTGGATCATCGGACCACCTGMAACCCACCTGTGAGTACAAAGTGCC 80515

QY 1599 CAGGTGAGTGGGCGTGGGCGTTGAGCCCTTCAGGCTTGATGATGCCATCTTGCCTATTG 1658

Db 80516 CAGGTGAGTGGGCGTGGGCGTTGAGCCCTTCAGGCTTGATGATGCCATCTTGCCTATTG 80575

QY 1659 TGTGGATATGCACACACAGGAGCATGCGCAGGTGTGTGGGACCTGTGTCTGTGA 80635

Db 80576 TGTGGATATGCACACACAGGAGCATGCGCAGGTGTGTGGGACCTGTGTCTGTGA 80635

QY 1719 CAAATGCCCTAGGCTGGGGAATAGCTTGTGTGCATCAGAGGAGCAGCAGCAGGATG 1778

Db 80636 CAAATGCCCTAGGCTGGGGAATAGCTTGTGTGCATCAGG ---CACAGCAGCAAGTGTG 80692

QY 1779 GCTCAGCAAAAACACACAGGGGAAAATGAGGGGCGATCAATCACTGCAAAATTTATT 1838

Db 80693 GCTCAGCAAAAACACACAGGGGAAAATGAGGGGCGATCAATCACTGCAAAATTTATT 1838

QY 1839 ATGAGGCTCCCGCTAAAAAAAAGAGGTCTCTTCTTGATGAAAGAGGAGAGAGGG 1898

Db 80739 ----- 80738

QY 1899 GTTTGTCTTATTAATTAAGGAGAGAGCCCGCTCAAAAAATTAAGGAGAGAGAGCC 1958

Db 80739 ----- 80738

QY 1959 AAGACCCCGTGGTGTGTGTTTCCAGGGGAGCTGGAACCTTTAGAGGAGCGCTGG 2018

Db 80739 ----- 80738

QY 2019 AGAACCGGTATTTCAAGCCCTTCGAGAGAAAAGAGCGCGCCCAAAAAATTCCTC 2078

Db 80739 ----- 80738

QY 2079 CCGGGCGATTAAGAAATGTGGGCTCTCAAAAAAGATGAAGAGAAACCGAGTTGATG 2138

Db 80739 ----- 80738

QY 2139 TGTGTATTTTTTAAACTCAGGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTGAG 2198

Db 80739 ----- 80761

QY 2199 CGCCTTCTAGAGAAACAAAGAGGCTTCCATTCTGAGAGAGAAACAGCAGAGAAACAG 2258

Db 80762 CGCCTTCTAGAGAAACAAAGAGGCTTCCATTCTGAGAGAGAAACAGCAGAGAAACAG 80821

QY 2259 GCAGATATCCTGTATTAATTTCAAGAGATPAAGGCTCTCTAGAAATATCAAGCAAGGT 2318

Db 80822 GCAGATATCCTGTATTAATTTCAAGAGATPAAGGCTCTCTAGAAATATCAAGCAAGGT 80881

QY 2319 GAGAGACAGAGACACCGGTGGCAGTGGGCTATTTCCAGGTTGATGTTGGAGAC 2378

Db 80882 GAGAGACAGAGACACCGGTGGCAGTGGGCTATTTCCAGGTTGATGTTGGAGAC 80941

QY 2379 ATCCTTTCTAAAGGAACCTGAGTGGGAAAGAAACATGACGATATCTCAGAGAGGCTT 2438

Db 80942 ATCCTTTCTAAAGGAACCTGAGTGGGAAAGAAACATGACGATATCTCAGAGAGGCTT 81001

QY 2439 CCTCCAGGACAGAAATCAGCAGGTGGAAAGCCCTGAGCAGCATTCACTAAACATCA 2498

Db 81002 CCTCCAGGACAGAAATCAGCAGGTGGAAAGCCCTGAGCAGCATTCACTAAACATCA 81061

QY 2499 TTGAGCAATCTACACGCTAGGTTCCATTATGGGAATGGGAATATGTTGGAGAGGG 2558

Db 81062 TTGAGCAATCTACACGCTAGGTTCCATTATGGGAATGGGAATATGTTGGAGAGGG 81121

QY	2559	CTGCGTGGTCCCTTCATCTACTTCTCACA	CTAGGGTGGTTGAGAGGCTTGGAGCTAAC	2618
Db	81122	CTGCGTGGTCCCTTCATCTACTTCTCACA	CTAGGGTGGTTGAGAGGCTTGGAGCTAAC	8118
QY	2619	AACAAGATGGGCTAGAAACACTGGCTTGG	CCAGAGGACCTGAGCTTAGTGTAAGACTT	2678
Db	81182	AACAAGATGGGCTAGAAACACTGGCTTGG	CCAGAGGACCTGAGCTTAGTGTAAGACTT	81241
QY	2679	GCTGCTGTACTGCTTGTGCTGTTATTTA	TTATTTATTTATTTATTTATGTCCTTAGA	2738
Db	81242	GCTGCTGTACTGCTTGTGCTGTTATTTA	TTATTTATTTATTTATTTATGTCCTTAGA	81301
QY	2739	GAGTTTGTCTTCTTATACCAAGGCTTGA	GGCGAAATGGCTATCTCAGCTCACTGCAM	2798
Db	81302	GAGTTTGTCTTCTTATACCAAGGCTTGA	GGCGAAATGGCTATCTCAGCTCACTGCAM	81361
QY	2799	TCCACCTCCTGGGATCAAGCGATTCTC	TGCCCCAGCTGAGAGCTTAGTGTAAGACTT	2858
Db	81362	TCCACCTCCTGGGATCAAGCGATTCTC	TGCCCCAGCTGAGAGCTTAGTGTAAGACTT	81421
QY	2859	GCACCCGACACACGCGCTGGATTAATTT	TTTGTATTTTATTTATTTATTTATTTAT	2918
Db	81422	GCACCCGACACACGCGCTGGATTAATTT	TTTGTATTTTATTTATTTATTTATTTAT	81481
QY	2919	GTTGGCCAGGCTGTGCTCGAATCCTCTG	ACCTTAGTGTATCCACGCGCTGACCTTCCA	2978
Db	81482	GTTGGCCAGGCTGTGCTCGAATCCTCTG	ACCTTAGTGTATCCACGCGCTGACCTTCCA	81541
QY	2979	ACTGCTGGGATTTATAGGATGAGCAGC	ATCGCCCACTGATTTATGAAAGTTAAAG	3038
Db	81542	ACTGCTGGGATTTATAGGATGAGCAGC	ATCGCCCACTGATTTATGAAAGTTAAAG	81601
QY	3039	TGGCAATGCACACGCGCTATCTACCTCT	TCCCTGCGCAAGAAAGGAGGAGGCTCTG	3098
Db	81602	TGGCAATGCACACGCGCTATCTACCTCT	TCCCTGCGCAAGAAAGGAGGAGGCTCTG	81661
QY	3099	ACTTCTTGGCTTCTACTTCCAAAGGAC	TGAGAACTGGAGGCGCTTGGAGACACT	3158
Db	81662	ACTTCTTGGCTTCTACTTCCAAAGGAC	TGAGAACTGGAGGCGCTTGGAGACACT	81721
QY	3159	TCATATCAGCTCTAGGGTCCCTTATGG	AGAGTTGAGGTCCAGACGAGGAGGGATCT	3218
Db	81722	TCATATCAGCTCTAGGGTCCCTTATGG	AGAGTTGAGGTCCAGACGAGGAGGGATCT	81781
QY	3219	CAGGCTGTGACACAGGGGCTCTGATCC	CTTACAAACCCCAATGGATGCCCTCTGA	3278
Db	81782	CAGGCTGTGACACAGGGGCTCTGATCC	CTTACAAACCCCAATGGATGCCCTCTGA	81841
QY	3279	GACCCAAAGCCACTCTCTCAGAGCCAC	TGCGCGGCAATGACATCACTTACAAAG	3338
Db	81842	GACCCAAAGCCACTCTCTCAGAGCCAC	TGCGCGGCAATGACATCACTTACAAAG	81901
QY	3339	GTTGGCTAAAGCCCGCTTAGGCTCTCT	TCCCGCTGCTGTGCGGGGAGGACATC	3398
Db	81902	GTTGGCTAAAGCCCGCTTAGGCTCTCT	TCCCGCTGCTGTGCGGGGAGGACATC	81961
QY	3399	TACAAGCTGTATATGGGAGATTCCTAT	CTTCCTCGCTCTGCTACTACATATCCG	3458
Db	81962	TACAAGCTGTATATGGGAGATTCCTAT	CTTCCTCGCTCTGCTACTACATATCCG	82021
QY	3459	ATTATTAAGTAAAGCTGGCAGAGGCTT	GGGCGCGGGGAGTGTGGCTGGGGGCT	3518
Db	82022	ATTATTAAGTAAAGCTGGCAGAGGCTT	GGGCGCGGGGAGTGTGGCTGGGGGCT	82078
QY	3519	GGGAGCTGGAGCTCTCTGGGGGCGCTCC	ACACCACTAGGAGCCAGTGCACCACTC	3578
Db	82079	GGGAGCTGGAGCTCTCTGGGGGCGCTCC	ACACCACTAGGAGCCAGTGCACCACTC	82134
QY	3579	ACAACACTAAAGCTGGGCTCTGACAC	AGCTCTGGGACCTGAGGCTGGCGGCTG	3638
Db	82135	ACAACACTAAAGCTGGGCTCTGACAC	AGCTCTGGGACCTGAGGCTGGCGGCTG	82193


```

misc_feature      /note="assembly_name:Config27"
30651. 35202
/note="assembly_name:Config28"
misc_feature      35303. 40201
/note="assembly_name:Config29"
misc_feature      40302. 45758
/note="assembly_name:Config30"
misc_feature      45859. 51249
/note="assembly_name:Config31"
misc_feature      51350. 58267
/note="assembly_name:Config32"
misc_feature      58368. 66916
/note="assembly_name:Config33"
misc_feature      67017. 77145
/note="assembly_name:Config34"
misc_feature      77246. 86493
/note="assembly_name:Config35"
misc_feature      86594. 96688
/note="assembly_name:Config36"
misc_feature      96789. 107169
/note="assembly_name:Config37"
misc_feature      107270. 117047
/note="assembly_name:Config38"
misc_feature      117148. 127120
/note="assembly_name:Config39
clone_end:SP6
vector_side:right"
misc_feature      127221. 142028
/note="assembly_name:Config40"
misc_feature      142130. 160169
/note="assembly_name:Config41"

```

Query Match	47.68;	Score 2379.8;	DB 2;	Length 160169;
Best Local Similarity	85.3%;	Pred. No. 0;		
Matches 2840;	Conservative 0;	Mismatches 122;	Indels 367;	Gaps 8;
QY 817 CATCTCTGTGAAGCACTTACTGATGGCCCTGGCCAGCCAAATCAGCCAGAGATAAAGTA				876
DB 4199 CACGNN				4258
QY 877 TGAATGTCACAGCAGCGCAATCAGAGCTCCTCGTCAGCATATGCAAGAAATTCGTCAATTTA				936
DB 4259 NNN				4318
QY 937 CTAGGGTGAATGAATATCCCAAGCAACACATCCTTTTCAATTAAGGCACTGAGGCTGAG				996
DB 4319 CTAGGGTGAATGAATATCCCAAGCAACACATCCTTTTCAATTAAGGCACTGAGGCTGAG				4378
QY 997 AAGAGAGCTGAACCTAACCCGGGGTACACACACAGGTGGCAAGGCTGGGACCAAGAAC				1056
DB 4379 AAGAGAGCTGAACCTAACCCGGGGTACACACACAGGTGGCAAGGCTGGGACCAAGAAC				4438
QY 1057 CAGGACTGTGACTGACACCCCGGATTCATCTTTCCATATAGCCACAGGGCTGTCAAAAG				1116
DB 4439 CAGGACTGTGACTGACACCCCGGATTCATCTTTCCATATAGCCACAGGG -TGTCAAAAG				1176
QY 1117 CCCCAGGGGCTAGTCAGAGGCTCCTCCTTCTGTGAGAGTTCTGTGGACAGAAAGTTGAAGC				4497
DB 4498 CCCCAGGGGCTAGTCAGAGGCTCCTCCTTCTGTGAGAGTTCTGTGGACAGAAAGTTGAAGC				4557
QY 1177 TCAGACAGGCCCCCTAAACCCCAACTCTCTCTGCAAGGGCTCAGGGGCTAGAAGCACTGGT				1236
DB 4558 TCAGACAGGCCCCCTAAACCCCAACTCTCTCTGCAAGGGCTCAGGGGCTAGAAGCACTGGT				4617
QY 1237 GGAGCAATATCTTTAGCCTCTGATTTTAGGGCATGTGAGAGGGGGGTGGTCCCTAAAT				1296
DB 4618 GGAGCAATATCTTTAGCCTCTGATTTTAGGGCATGTGAGAGGGGGGTGGTCCCTAAAT				4677
QY 1297 TCCAGCCCTGGTCTCAGCCCAACACCCCTCAAGAAATAAGAGGGGCTATGGCCAGGC				1356
DB 4678 TCCAGCCCTGGTCTCAGCCCAACACCCCTCAAGAAATAAGAGGGGCTATGG -CAGGC				4736

QY	1357	TGTCCTAGCCGTCCTCTCTGAGCAGATTACAAAGGAGACTAAGACAAAGACTCTTGT	1416
Db	4737	TGTCCTAGCCGTCCTCTCTGAGCAGATTACAAAGGAGACTAAGACAAAGACTCTTGT	4796
QY	1417	GGAGGTCCTGAGCCTTAGGGAGTCAAGTACGCGGCTCAGCCTCAGCGGGCAGTGGCAG	1476
Db	4197	GGAGGTCCTGAGCCTTAGGGAGTCAAGTACGCGGCTCAGCCTCAGCGGGCAGTGGCAG	4856
QY	1477	CCTCTAAGAGTGGGACAGGGGCACAGAGTCCAGGAGTCCACAGGCTTAGTC	1536
Db	4857	CCTCTAAGAGTGGGACAGGGGCACAGAGTCCAGGAGTCCACAGGCTTAGTC	4916
QY	1537	GCCAGACCTCTGTGGGATCATGGGACCACCTGAAACCCACCTGTGATCAAGGTGC	1596
Db	4917	GCCAGACCTCTGTGGGATCATGGGACCACCTGAAACCCACCTGTGATCAAGGTGC	4976
QY	1597	CCGAGTGGAGTGGGCTGGGGCTTTGAGGCCCTTCAGGGTTGGATGGGCATCTGCTATT	1656
Db	4977	CCGAGTGGAGTGGGCTGGGGCTTTGAGGCCCTTCAGGGTTGGATGGGCATCTGCTATT	5036
QY	1657	TGTCGGGATATGCACACAGGACACATGCGCAGGTGTGTGGGACCTGTGTCTG	1716
Db	5037	TGTCGGGATATGCACACAGGACACATGCGCAGGTGTGTGGGACCTGTGTCTG	5096
QY	1717	TGCAAAAGCCCTGAGGTGGGAATAGTGTGTGTGCAATCAGAGACGACGACGACCCAGTG	1776
Db	5097	TGCAAAAGCCCTGAGGTGGGAATAGTGTGTGTGCAATCAG--CAAGGACACCCAGTG	5153
QY	1777	TGCGTCGACGACAAACACAGGAGGAAAGATGAGAGGGGACATCAATCACTGCACAAATAT	1836
Db	5154	TGCGTCGACGACAAACACAGGAGGAAAGATGAGAGGGGACATCAATCACTGCACAAATAT	5201
QY	1837	TTATAGAGCTCCCTTAAAAAAGAAAGTCTCTTTGATAGAAAGAGGAGAGG	1896
Db	5202	-----	5201
QY	1897	GGGTGTCTTATTAATATTAAGGAGAGCCGCCCTCAAAAAATTAAGGAGGAGAGAC	1956
Db	5202	-----	5201
QY	1957	CCAAAGACCCGTGGGTGTGTGTTTCCAGGGGAGCTCGAACCTTTAGAGGAGCGTG	2016
Db	5202	-----	5201
QY	2017	GGAGAACCGCTGTATTTCAGGCCCTCCAGAGAAAAAGAGCGGCCCCCAAAAAATATCCC	2076
Db	5202	-----	5201
QY	2077	TCGGGGGAGATAAAGAAATGCTGGGCTCTCTCAAAAAAGTGAAGAGAAAGCCGGATTGTA	2136
Db	5202	-----	5201
QY	2137	TGCTGTGATATTTTAAAACTCAGGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTG	2196
Db	5202	-----	5222
QY	2197	AGCGCCTTCTACGAGAACACAAAGAGAGCTTCATTCGAGGAGAAACAGGACGAAC	2256
Db	5223	AGCGCCTTCTACGAGAACACAAAGAGAGCTTCATTCGAGGAGAAACAGGACGAAC	5282
QY	2257	AGGAGATATCTGTATTAATTCAAGTGTGATTAAGTGCCTCTAGAAATATCAAGCAAG	2316
Db	5283	AGGAGATATCTGTATTAATTCAAGTGTGATTAAGTGCCTCTAGAAATATCAAGCAAG	5342
QY	2317	GTGAGGAGACACAGACCCGGTGGCAGTGGGCTTATTTCAGGTTGATGGTGGGA	2376
Db	5343	GTGAGGAGACACAGACCCGGTGGCAGTGGGCTTATTTCAGGTTGATGGTGGGA	5402
QY	2377	ACATTCCTTTCAAAGGGAACCTGGAGTGGGAAAGAACCTCAGATATCTCAGAAAGAC	2436
Db	5403	ACATTCCTTTCAAAGGGAACCTGGAGTGGGAAAGAACCTCAGATATCTCAGAAAGAC	5462

QY	2437	TTCTCCAGAGGAAAGATACAGAGGAGGAAAGGCCCTGGAGGCCACATTCAGTAAACAT	2496
Db	5463	TTCTCCAGGAGGAAAGATCAGCAGGTGGAAAGGCCCTGGAGGCCACATTCAGTAAACAT	5522
QY	2497	CATTGAGCACTCTTACAGAGCTAGTTCACATTATGAGGAATGGGAATATGCTGTGACAG	2556
Db	5523	CATTGAGCACTCTTACAGAGCTAGTTCACATTATGAGGAATGGGAATATGCTGTGACAG	5582
QY	2557	GGCTGGCTGGTCCCTTCCATATCTTCTCAGACTAGGGTGTGGAGAGACTTGGGAGCTAA	2616
Db	5583	GGCTGGCTGGTCCCTTCCATCTCTTCTCAGACTAGGGTGTGGAGAGACTTGGGAGCTAA	5642
QY	2617	CGAACCAAGATGGGTGAGAACACCTGGAGCCAGAGAGACCTGTAGCTTAGCTTAAGCA	2676
Db	5643	CGAACCAAGATGGGTGAGAACACCTGGAGCCAGAGAGACCTGTAGCTTAGCTTAAGCA	5702
QY	2677	TTGCTGCTGTACTGCTTGTGCTGCTGTGATTTATTTATTTATTTATTTATTTATTTATTTA	2736
Db	5703	TTGCTGCTGTACTGCTTGTGCTGCTGTGATTTATTTATTTATTTATTTATTTATTTATTTA	5762
QY	2737	CAGGTTTGTGCTCTTACACAGGCTGAGTGAATGGGTGATATCAGCTACCTGCA	2796
Db	5763	CAGGTTTGTGCTCTTACACAGGCTGAGTGAATGGGTGATATCAGCTACCTGCA	5822
QY	2797	CCTCCACCTCCTGGGATCAAGCGATTCTCTGCTCAGCCTCCTGAGTAGCTGGGATTA	2856
Db	5823	CCTCCACCTCCTGGGATCAAGCGATTCTCTGCTCAGCCTCCTGAGTAGCTGGGATTA	5882
QY	2857	AGGACACCGGACACAGGCTGATTAATTTTTTTTATTTATTTATTTATTTATTTATTTATTTA	2916
Db	5883	AGGACACCGGACACAGGCTGATTAATTTTTTTTATTTATTTATTTATTTATTTATTTATTTA	5942
QY	2917	ATGTTGGCGACAGCTGAGTCTCGAACCTCGAACCTTAGTGATCCACTGCTCGACTTCCC	2976
Db	5943	ATGTTGGCGACAGCTGAGTCTCGAACCTCGAACCTTAGTGATCCACTGCTCGACTTCCC	6002
QY	2977	AAAGTGTGGGATTATAGGCATAGGCACACTGCGCCAGTATTTATGAAAGTTTAAAGCA	3036
Db	6003	AAAGTGTGGGATTATAGGCATAGGCACACTGCGCCAGTATTTATGAAAGTTTAAAGCA	6062
QY	3037	CATGGCAATGACACAGGCTATCTAGCTTCCCTGCGCAAAAGGAGGAGCCTCTG6GC	3096
Db	6063	CATGGCAATGACACAGGCTATCTAGCTTCCCTGCGCAAAAGGAGGAGCCTCTG6GC	6122
QY	3097	TCACCTTCTTGCGTTTCTACTTCCAAAAGGAGTCAAGACTGCGAGGGCCTTGGAGACCA	3156
Db	6123	TCACCTTCTTGCGTTTCTACTTCCAAAAGGAGTCAAGACTGCGAGGGCCTTGGAGACCA	6182
QY	3157	CTTCAATCCACTCCTAAGGTCGCCATGSGAAGTGAAGTGCAGAGCAGGGAAAGGTCCT	3216
Db	6183	CTTCAATCCACTCCTAAGGTCGCCATGSGAAGTGAAGTGCAGAGCAGGGAAAGGTCCT	6242
QY	3217	GACAGGCTCTACCAAGGGCTCTATATCCTTCCAAAACCCCAATGGGTGCCCTCTTAC	3276
Db	6243	GACAGGCTCTACCAAGGGCTCTATATCCTTCCAAAACCCCAATGGGTGCCCTCTTAC	6302
QY	3277	AGGACCCAGGCCACCTGCTGACAGCCACTGCTGGCCATACATCATCTTACACAAC	3336
Db	6303	AGGACCCAGGCCACCTGCTGACAGCCACTGCTGGCCATACATCATCTTACACAAC	6366
QY	3337	AAGTAGCTAAATGAGCCGTTAGGCTCTCTCTCCGCTGCTGCTGTGCTGCTGCTGCTGCTG	3396
Db	6363	AAGTAGCTAAATGAGCCGTTAGGCTCTCTCTCCGCTGCTGCTGTGCTGCTGCTGCTGCTG	6422
QY	3397	TCATCAAGCTCTATATGCGGATTCCTATCTCTGCTCTGCTACTACATCATCCCT	3456
Db	6423	TCATCAAGCTCTATATGCGGATTCCTATCTCTGCTCTGCTACTACATCATCCCT	6482
QY	3457	TTATTTATAGTAAAGCTGCGAGGGGCTGGCGGGGGGCTGGGAAAGATGTGGCTGGG	3516
Db	6483	TTATTTATAGTAAAGCTGCGAGGGGCTGGCGGGGGGCTGGGAAAGATGTGGCTGGG	6541
QY	3517	CTGGAGCTGGGAGCTCCTGGGGGCTTCCACAGCAAGCTCAGGGGCCAGTGCACAGTCCA	3576

Db 6542 -----GCTGGAGACTCCTGGGGGCTGCCAGCAGCTCAGGGGCTGTACAGTCA 6595
 QY 3577 CTACACAGCTAGTGGGCTCTCTACAGCTCTCTGGGCTAGTGGAGCTGGGCT 3636
 Db 6596 CTACACAGCTAGTGGGCTCTCTACAGCTCTCTGGGCTAGTGGAGCTGGGCT 3636
 QY 3637 GGGGGCTGGGCTAGTAAAGTCACTAGTGGGCTCTCTGGGCTAGTGGAGCT 6654
 Db 6655 GGGGGCTGGGCTAGTAAAGTCACTAGTGGGCTCTCTGGGCTAGTGGAGCT 6654
 QY 3697 TTAGGACCTCTCTCTCAACCCAGAGAGAGCTGGAGCTGGAGCTGGAGCTGG 3756
 Db 6715 TTAGGACCTCTCTCTCAACCCAGAGAGAGCTGGAGCTGGAGCTGGAGCT 6715
 QY 3757 ACTCTAGCTGGCTTTTCTCTCAACCCAGAGAGAGCTGGAGCTGGAGCTGG 6774
 Db 6775 ACTCACTGGCTTTTCTCTCAACCCAGAGAGAGCTGGAGCTGGAGCTGGAG 6834
 QY 3817 AAA--AG 6834
 Db 6835 AAAAG 6894
 QY 3875 GGTCTGAACCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6894
 Db 6895 GGTCTGAACCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6954
 QY 3935 ACAGGTGAGCTCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 6954
 Db 6955 ACAGGTGAGCTCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 6954
 QY 3994 TTGAG 7014
 Db 7015 TTGAG 7014
 QY 4054 ACTGTAACCTCTCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 7074
 Db 7075 ACTGTAACCTCTCTCTCTCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 7074
 QY 4114 GAATTATAGGAC 4142
 Db 7135 GAATTATAGGAC 7163

RESULT 8
 LOCUS AX347149/c 6396 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 2220 from Patent WO200928.
 ACCESSION AX347149
 VERSION AX347149.1 GI:18495037
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1
 AUTHORS Olek A., Piepenbrock, C. and Berlin, K.
 TITLE Diagnosis of diseases associated with the immune system
 JOURNAL Patent: WO 0200928-A 2220 03-JAN-2002;
 FEATURES
 source Location/Qualifiers
 1. 6396
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 1570 a 89 c 1585 g 3152 t
 ORIGIN

Query Match 25.88: Score 1291.8; DB 6; Length 6396;
 Best Local Similarity 66.98; P-Val. No. 0;
 Matches 2111; Conservative 0; Mismatches 687; Indels 357; Gaps 5;
 123 CAACCCACATGACAGTATACCTATGTAACCAACCTGACCATATATATACCTATG 182

Db 2800 CAACCCACATGACAGTATACCTATGTAACCAACCTGAC-CATCATATATACCTATA 2742
 QY 183 TAACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
 Db 2741 TAACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
 QY 243 TGTGTAG 302
 Db 2683 TAT 2624
 QY 303 AGGGTGAAG 362
 Db 2623 AAAT 2564
 QY 363 AGGGTGAAG 422
 Db 2563 AAAT 2504
 QY 423 TAAATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2504
 Db 2503 TAAATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2444
 QY 483 CCGTTTAAAT 2444
 Db 2443 CCGTTTAAAT 2384
 QY 543 TAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 Db 2383 TAAATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2324
 QY 603 TTCAACAGCTTAG 662
 Db 2223 TTCAACAGCTTAG 2264
 QY 663 ATGCTGAGATTTTACAG 722
 Db 2263 ATGCTGAGATTTTACAG 2204
 QY 723 GCTGGCTGGAAG 782
 Db 2203 ACTTAT 2144
 QY 783 CCCCCCTGACAG 842
 Db 2143 CCCCCCTGACAG 2084
 QY 843 GGGCCCTGACAG 902
 Db 2083 AACCTTACAG 2024
 QY 903 TCCCTGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
 Db 2023 TCCCTGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1964
 QY 963 AACATCTCTTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
 Db 1963 AACATCTCTTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 QY 1023 AACACACAG 1082
 Db 1903 AACACACAG 1844
 QY 1083 TCATCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
 Db 1843 TCATCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
 QY 1143 CTCTCTGAG 1202
 Db 1783 CTCTCTGAG 1724
 QY 1203 CTCTCTGAG 1262
 Db 1723 CTCTCTGAG 1664

1263 TTAGGGCCATGTAGAGGGGGTGTGCTTAAATTCAGCCCTGGTCTGACCCACACC 1322
1663 TTAAAAACCATTAATAAAAAATATATACCTTAATTCACACCTTAATCTCAACCAACACC 1604
1323 CTCGCAAGAAAGAAATTAAGAGGGGCGATGGCGCTGGAGCGGTGCTCTGAGCAGA 1382
1603 CTCGCAAAAAAATTAATAAAAAACCTTAACCAAACTTACTACCGGTACTCTTCAACAA 1544
1383 TTCAAGAAAGGAACTAAGACAGAGACTCTTGTGAGAGTCTGTGGATGAGGATCAAGT 1442
1543 TTCAAAAAAATTAATAAAAAACCTTAAATTAATTAATTAATTAATTAATTAATTAAT 1484
1443 GAGGGGGCTGACGACGCTGAGGGGAGTGGCCAGCCTCTAAGAGTGGCAGGGGCACTGG 1502
1483 AAGGAGACTCAACGACTCAAGTAAACAACTTAATTAATTAATTAATTAATTAATTAAT 1424
1503 CCACAGAGTCCAGGAGATCCACAGCCTAGTGGCAGACCTTGTGGGATCATCGGA 1562
1423 CCACAAATATCCCAAAAAATCCCAACCTTAATCCGCAAACTCTTAATTAATTAATTAAT 1364
1563 CCACCTGGAACCCCACTGCTGAGTCAAGAGTGGCCAGGAGTGGAGTGGGCTGGTGG 1622
1363 CCACCTAAACCCCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1304
1623 AGGCTTCAAGGTTGGATGGCATCTTGGTATTTGGTGGATGACACACAGCAG 1682
1303 AAACCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1244
1683 CACATCCGACAGTGTGTGGCAGCCTGTGTCTGTGCAATGCGCTGAGGTGGGATGAG 1742
1243 CACATACGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1184
1743 CTGTGTGTCATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1802
1183 CTAAATATATCATC---AATACACACCAACCAATATATACATCAACCAACCAACCA 1127
1803 GAATGAGGGGGGCTCAATCACTGACAAATTAATTAATTAATTAATTAATTAATTAAT 1862
1126 AATATAAAAAACAATCAATCAC----- 1105
1863 AGGTCTCTTCTTGCATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1922
1104 ----- 1105
1923 GGAGCCGCCCTCAAAAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982
1104 ----- 1105
1983 CCAGGGGAGCTCGAACCTTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2042
1104 ----- 1105
2043 GAGAGAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2102
1104 ----- 1105
2103 CTCTCAAAAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2162
1104 ----- 1105
2163 TAGNNNNNNNNNGCTCAGTAATTTTATATGAGGCGCTTCTACGAGAACACAGAGG 2222
1104 -----TACTTCATTAATTTTATTAAGCGCTTCTACGAGAACACAGAGG 1058
2223 AGCTTCATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2282
1057 AACTTCATCTCAAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 998
2283 TAGGATTAAGTCTCTGAGAAATATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2342
997 TAAATATTAATTAATCTTAATAAATATCAACAACAACAACAACAACAACAACAACAACA 938

QY 2343 AGTGGGCTCTATTTCCAGGTTGATGTTGGGAACATCTTCTAAAGGAGACCTGGAG 2402
DB 937 AATTAACCTCTATTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 878
QY 2403 TGGGAAGAACCATGACGATATCTCAGAGAGGCTTCTCCAGGACGAGAGATCAGCAGG 2462
DB 877 TAAAAAACCATACCAATATCTCAAAAAAATCTTCCTCCCAAAAAAATTAATCAACAA 818
QY 2463 TGGAAAGGCCCTGGAGGACCATCTGCTAATCATCTTGGAGATCTACAGGATAGGT 2522
DB 817 TAAAAAACCCTTAACCAACCACTTCAATTAACCACTTAAACATCTTAACCACTTAAT 758
QY 2523 TCCATTAATGGAATGGGAATATGTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2582
DB 757 TCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 698
QY 2583 CACACTAGGAGGTGAGAGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 2642
DB 697 CACACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 638
QY 2643 CTAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2702
DB 637 CTAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 578
QY 2703 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2762
DB 577 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 518
QY 2763 TTGAGTGCATAGGCTGATCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 2822
DB 517 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 458
QY 2823 CTCGCTGCTGACGCTGCTGAGTGTGGATTAAGGAGGAGGAGGAGGAGGAGGAGGAG 2882
DB 457 CTCCTTACCTCAACCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 398
QY 2883 TTTTGTATTTTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2942
DB 397 TTTTGTATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 338
QY 2943 CTGACCTTAAGTATGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3002
DB 337 CTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 278
QY 3003 CACGCGGCGGAGTGTATTAAGAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3062
DB 277 CACTAGGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 218
QY 3063 TCTTCCCTGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3122
DB 217 TCTTCCCTGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 158
QY 3123 AAGGAGTCAAGACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3182
DB 157 AAAAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 98
QY 3183 GAGAGGAGTGAAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3242
DB 97 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 38
QY 3243 CCCTACAAACCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330
DB 37 CCCTACAAACCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3

RESULT 9
AX347148
LOCUS AX347148 6396 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 2219 from Patent WO0200928.
ACCESSION AX347148
VERSION AX347148.1 GI:18495036
KEYWORDS
SOURCE synthetic construct.

2042 CGAGAGAAAGAGAGGCGCCGCCAAAAATATCCCTCCCGGCGATTAAGAAATGGTGCC 2101
5292 ----- 5291
2102 TCTCTCAAAAGATGAGAGAGAGCGGAGTTGTATGTGTATTTTAAACTCCAG 2161
5292 ----- 5291
2162 GTAGNN 2221
5292 -----TTGTTTGTAGTAAATTTTATTGAGCGTTTACAGAGATATTAAG 5338
2222 GAGCTTCATCTGAG 2281
5339 GAGTTTATTTTGTAG 5398
2282 GTAGTATAGTGTCTCTAGAAATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2341
5399 GTAGTATAGTGTCTCTAGAAATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5458
2342 CAGTGGGCTCTATTTCCAGGTGGATGTTGGGAACATCTTCTTAAAGGAACTGGA 2401
5459 TAGTGGGCTCTATTTTCCAGGTGGATGTTGGGAACATCTTCTTAAAGGAACTGGA 5518
2402 GTGGAG 2461
5519 GTGGAG 5578
2462 GTGGAG 2521
5579 GTGGAG 5638
2522 TTTCATATAG 2581
5639 TTTCATATAG 5698
2582 TCACACTAG 2641
5699 TTATATAG 5758
2642 CCTGAGCCAG 2701
5759 TTGAGTATAG 5818
2702 TGTAT 2761
5819 TGTAT 5878
2762 CTGAGTGCAG 2821
5879 TTGAGTGCAG 5938
2822 TCTGCTGCTGAG 2881
5939 TTTTAT 5998
2882 TTTTAT 2941
5999 TTTTAT 6058
2942 CCTGAGCTGAG 3001
6059 TTTGAT 6118
3002 CCACTGGGCGCAGAT 3061
6119 TTTATGCGTTTAT 6178
3062 GTCTTCCCTGCAAG 3121
6179 GTTTTAT 6238

3122 AAAGCACTGAACTGCGAG 3181
6239 AAAGTATGTTAGATTTGATAG 6298
3182 TGGGAGAGTGTAGAGTCCGAG 3241
6299 TGGGAGAGTGTAGAGTCCGAG 6358
3242 TCCCTACAAACCCCAATCGGTGTCCTCTACAGG 3279
6359 TTTTAT 6396
RESULT 10
AC051664/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-810P12, WORKING DRAFT
AC051664
AC051664.2 GI:7637836
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens.
REFERENCE
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 160169)
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
JOURNAL
2 (bases 1 to 160169)
Waterston, R.H.
Submitted (15-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 23, 2000 this sequence version replaced gi:7574984.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project information
Center project name: H.NH0810P12
----- Summary Statistics -----
Sequencing vector: plasmid; 100%
Chemistry: Dye-terminator; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 147316 bases at least Q40
Consensus quality: 151042 bases at least Q30
Consensus quality: 152875 bases at least Q20
Insert size: 16500; agarose-fp
Insert size: 157869; sum-of-coverage
Quality coverage: 3.91 in Q20 bases; sum-of-coverage
Quality coverage: 4.17 in Q20 bases; sum-of-coverage
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1926: contig of 1926 bp in length
1927 2026: gap of unknown length
2027 4202: contig of 2176 bp in length
4203 4302: gap of unknown length
4303 7163: contig of 2861 bp in length
7164 7263: gap of unknown length
7264 9961: contig of 2698 bp in length
9962 10061: gap of unknown length


```
* 10062 12575: contig of 2514 bp in length
* 12576 12575: gap of unknown length
* 12675 15545: contig of 2870 bp in length
* 15546 15645: gap of unknown length
* 15646 20941: contig of 5296 bp in length
* 20942 21041: gap of unknown length
* 21042 23966: contig of 2925 bp in length
* 23967 24066: gap of unknown length
* 24067 27231: contig of 3165 bp in length
* 27232 27332 30530: contig of 3219 bp in length
* 30531 30650: gap of unknown length
* 30651 35202: contig of 4552 bp in length
* 35203 35303 40201: contig of 4899 bp in length
* 40202 40301: gap of unknown length
* 40302 45758: contig of 5457 bp in length
* 45759 45858: gap of unknown length
* 45859 51249: contig of 5391 bp in length
* 51250 51349: gap of unknown length
* 51350 58267: contig of 6918 bp in length
* 58268 58368 66916: gap of unknown length
* 58369 66916: contig of 8549 bp in length
* 66917 77016: gap of unknown length
* 77017 77145: contig of 10129 bp in length
* 77146 77245: gap of unknown length
* 77246 86493: contig of 9248 bp in length
* 86494 86594 96688: gap of unknown length
* 86594 96688: contig of 10095 bp in length
* 96689 96788: gap of unknown length
* 96789 107169: contig of 10381 bp in length
* 107170 107269: gap of unknown length
* 107270 117047: contig of 9778 bp in length
* 117048 117147: gap of unknown length
* 117148 127120: contig of 9973 bp in length
* 127121 142029: contig of 14809 bp in length
* 142030 142129: gap of unknown length
* 142130 160169: contig of 18040 bp in length.
FEATURES
source
1..160169 Location/Qualifiers
1..160169 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-810P12"
1..1926
/misc_feature /note="assembly_name:Contig18"
2027..4202
/misc_feature /note="assembly_name:Contig19"
4303..7163
/misc_feature /note="assembly_name:Contig20"
vector_end:r7
vector_side:left"
7264..9961
/misc_feature /note="assembly_name:Contig21"
10062..12575
/misc_feature /note="assembly_name:Contig22"
12676..15545
/misc_feature /note="assembly_name:Contig23"
15646..20941
/misc_feature /note="assembly_name:Contig24"
21042..23966
/misc_feature /note="assembly_name:Contig25"
24067..27231
/misc_feature /note="assembly_name:Contig26"
27332..30530
/misc_feature /note="assembly_name:Contig27"
30651..35202
/misc_feature /note="assembly_name:Contig28"
35303..40201
/misc_feature /note="assembly_name:Contig29"
40302..45758
/misc_feature /note="assembly_name:Contig30"
```

```
misc_feature 45859..51249
/note="assembly_name:Contig31"
misc_feature 51350..58267
/note="assembly_name:Contig32"
misc_feature 58368..66916
/note="assembly_name:Contig33"
misc_feature 67017..77145
/note="assembly_name:Contig34"
misc_feature 77246..86493
/note="assembly_name:Contig35"
misc_feature 86594..96688
/note="assembly_name:Contig36"
misc_feature 96789..107169
/note="assembly_name:Contig37"
misc_feature 107270..117047
/note="assembly_name:Contig38"
misc_feature 117148..127120
/note="assembly_name:Contig39"
vector_end:SP6
vector_side:right"
misc_feature 127221..142029
/note="assembly_name:Contig40"
misc_feature 142130..160169
/note="assembly_name:Contig41"
BASE COUNT 41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN
Query Match 14.8% Score 742.2; DB 2; Length 160169;
Best Local Similarity 86.5% Pred. No. 5.6e-195;
Matches 776; Conservative 0; Mismatches 118; Indels 3; Gaps 2;
QY 4104 TGAGTACCTGGAATTATAGCAGACACACCAGCCCTGGCTAATTTTTTTTTTCTGT 4163
|
DB 77246 TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 77187
QY 4164 ATTTTAGTAGAGACAGAGGTTTCATCATGTGGCCAGCTGTCCTGAGACCCCTGACCTC 4223
DB 77186 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 77127
QY 4224 AATGTATCCACCCACCTCGGCTCCCAAGTGTGGATTACAGGTGTCACCCACATGC 4283
DB 77126 AATGTATCCACCCACCTCGGCTCCCAATGTGTCGAGTACAGGTGTCACCCATGC 77067
QY 4284 ACAGCCACATGATGATTTTAAATATTTTAAATATTTTAAATATTTTAAATTTTAA 4343
DB 77066 ACAGCCACATGATGATTTTAAATATTTTAAATATTTTAAATATTTTAAATTTTAA 77007
QY 4344 GTAGCAGTACCTGCTGTATATCCAGACTTTGAGGGCCAGGTGGGGATCACT 4403
DB 77006 GTAGCAGTACCTGCTGTATATCCAGACTTTGAGGGCCAGGTGGGGATCACT 76947
QY 4404 TGAGCTGGAGTTACGCTGGGCAACATAGTAGACCCCGTCTACCAAAATTTTAA 4463
DB 76946 TGAGCTGGAGTTACGCTGGGCAACATAGTAGACCCCGTCTACCAAAATTTTAA 76887
QY 4464 AATTAGCTGGAGTGGTGGCATTTGCTGCTCCCACTACTTGGAGAGTGT 4523
DB 76886 AATTAGCTGGAGTGGTGGCATTTGCTGCTCCCACTACTTGGAGAGTGT 76827
QY 4524 GGGATGGCTGAAGCTGTGAGGTGAGGCTCAGTACCTATGATCAACACCTGACT 4583
DB 76826 GGGAT -GCTGAAGCTGTGAGGTGAGGCTCAGTACCTATGATCAACACCTGACT 76768
QY 4584 TCAGCTGAGTACAGGCTATCTCAAAAGCAAAATATATGTTATCTAAACGTAG 4643
DB 76767 TCAGCTGAGTACAGGCTATCTCAAAAGCAAAATATATGTTATCTAAACATANA 76708
QY 4644 GTATATACAGAAATATATGATAGCATTTTAAATGAAGAAACATTAATGATTAATGGA 4703
DB 76707 ATATATACAGAAATATATGATAGCATTTTAAATGAAGAAACATTAATGATTAATGGA 76648
QY 4704 TTGTAATAATCAATATCATGAATTTCTGTGTTATATATGTCAGCAACAGGACA 4763
|||||
```



```

Db      76647 TTGTAATAATATACAAATGAAATCTTGCTGTTCTTAATATATGCTAGCAACAGCACACA 76588
Qy      47664 TTTGCTTTTACTAGGACACCAAGTACTTTTAAAAAAGTTAGGCGCACAGGCGCT 4823
Db      76587 TTTGGTTTTTACTAGGGACCAAGGTACTTTAAAAAAGTTAGGCGCACAGGCGCT 76528
Qy      4824 CACACCTGTATATCCCGACACTTTGGAGGCGCAAGGAGATCATCTTGTAGCCGAGAG 4883
Db      76527 CACACCTGTATATCCCGACACTTTGGAGGCGCAAGGAGATCATCTTGTAGCCGAGAG 76468
Qy      4884 TTTAGACCTGTAGCAACATAGGAGATCTGTATCTTATATATATATATATATAT 4943
Db      76467 TTTAGACCTGTAGCAACATAGGAGATCTGTATCTTATATATATATATATATAT 76408
Qy      4944 GCGTAGGCGCTTTGGCTTACACCCGTATATCCCGACACTTTGGAGGCGCGAGGCGGT 5000
Db      76407 G--CTAGGCCCTTGTCTTACACCCGTATATCCCGACACTTTGGAGGCGCGAGGCGGT 76353

```

```

RESULT 11
AC084857/c 133683 bp DNA linear HTG 22-NOV-2000
LOCUS Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT
DEFINITION
ACCESSION AC084857
VERSION AC084857
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abrahams, H., Allen, N.,
JOURNAL 1 (bases 1 to 133683)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Homo sapiens chromosome 11, clone CTD-3231N5
Unpublished
2 (bases 1 to 133683)

```

```

TITLE
JOURNAL
COMMENT
Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W1BR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: l11038
Center clone name: 3231_N5
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

```

FEATURES

source

```

misc_feature 1..1459 contig of 1459 bp in length
misc_feature 1560..2520 contig of 961 bp in length
misc_feature 2521..2620 contig of 100 bp
misc_feature 2621..6334 contig of 3714 bp in length
misc_feature 6335..6435 contig of 100 bp
misc_feature 6436..11913 contig of 5479 bp in length
misc_feature 11914..12013 contig of 100 bp
misc_feature 12014..20756 contig of 8743 bp in length
misc_feature 20757..20856 contig of 100 bp
misc_feature 20857..29735 contig of 8879 bp in length
misc_feature 29736..29835 contig of 100 bp
misc_feature 29836..38645 contig of 8810 bp in length
misc_feature 38646..38745 contig of 100 bp
misc_feature 38746..52193 contig of 13448 bp in length
misc_feature 52194..52293 contig of 100 bp
misc_feature 52294..65209 contig of 12916 bp in length
misc_feature 65210..65309 contig of 100 bp
misc_feature 65310..79796 contig of 14487 bp in length
misc_feature 79797..79896 contig of 100 bp
misc_feature 79897..95896 contig of 16000 bp in length
misc_feature 95897..113419 contig of 17423 bp in length
misc_feature 113420..113520 contig of 100 bp
misc_feature 113520..133683 contig of 20164 bp in length.
Location/Qualifiers
1..133683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-3231N5"
/clone_id="CITD Human BAC"
1..1459
/note="assembly-fragment"
1560..2520
/note="assembly-fragment"
2521..6334
/note="assembly-fragment"
6335..11913
/note="assembly-fragment"
12014..20756
/note="assembly-fragment"
20857..29735
/note="assembly-fragment"
29836..38645
/note="assembly-fragment"
38746..52193
/note="assembly-fragment"
52294..65209
/note="assembly-fragment"
65310..79796
/note="assembly-fragment"
79897..95896
/note="assembly-fragment"
95897..113419
/note="assembly-fragment"

```

Assembly program: Phrap; version 0.960731
Consensus quality: 126761 bases at least Q40
Consensus quality: 129952 bases at least Q30
Consensus quality: 131351 bases at least Q20
Insert size: 128000; agarose-fp
Insert size: 132483; sum-of-contigs
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

misc_feature /note="assembly_fragment"
 113520..113683
 BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
 ORIGIN

Query Match 14.6%; Score 728.2; DB 2; Length 133683;
 Best Local Similarity 87.3%; Pred. No. 4.3e-191;
 Matches 766; Conservative 0; Mismatches 108; Indels 3; Gaps 3;

QY 122 GCAACCCACCATGGACAGATATACCTATGTATACCACTGCACCATCATGTATACCTAT 181
 DB 7203 GCAACCCACCATGGACAGATATACCTATGTATACCACTGCACCATCATGTATACCTAT 7145
 QY 182 GTACCAACCTGTATACCTATGTATACCTATGTATACCACTGCACCATCATGTATACCTAT 241
 DB 7144 GTACCAACCTGTATACCTATGTATACCTATGTATACCACTGCACCATCATGTATACCTAT 7087
 QY 242 GTGTGTAGAAAAATCACCTGCATATCTACAGATATACCTATGTATACCTATGTATACCTAT 301
 DB 7086 GTGTGTAGAAAAATCACCTGCATATCTACAGATATACCTATGTATACCTATGTATACCTAT 7027
 QY 302 GAGGGTACAG 361
 DB 7026 GAGGGTACAG 6967
 QY 362 CAGGGCTTGTGAGAGTTTACCTCCAAACACACATCTCATATCCCTGCAGTGTGA 421
 DB 6966 CAGGGCTTGTGAGAGTTTACCTCCAAACACACATCTCATATCCCTGCAGTGTGA 6907
 QY 422 TTTAAATGCAATCCCTTAAGGCCACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
 DB 6906 TTTAAATGCAATCCCTTAAGGCCACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6847
 QY 482 CCGGTTTAAATGCAATCCCTTAAGGCCACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
 DB 6846 CCGGTTTAAATGCAATCCCTTAAGGCCACAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6787
 QY 542 CTAGAGCTGTGCTCTCTTATTTGAACAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 DB 6786 CTAGAGCTGTGCTCTCTTATTTGAACAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6727
 QY 602 CTTCACACAGCTTAG 661
 DB 6726 CTTCACACAGCTTAG 6667
 QY 662 CATGCTGAGATTTTACACAGCTGAG 721
 DB 6666 CATGCTGAGATTTTACACAGCTGAG 6607
 QY 722 GCGTGTGAG 781
 DB 6606 GCGTGTGAG 6547
 QY 782 ACCCCCTTCACACAG 841
 DB 6546 ACCCCCTTCACACAG 6487
 QY 842 GGGGCTTCCAG 901
 DB 6486 GGGGCTTCCAG 6427
 QY 902 CTCTGCTGAGCATATGAGAGATTTCTGATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
 DB 6426 NNN 6367
 QY 962 CACCATCTTTTACAGATTAAGGAG 998
 DB 6366 NNN 6330

RESULT 12
 CINS07ECV

LOCUS CINS07ECV 165423 bp DNA linear PRI 22-AUG-2001
 DEFINITION Human chromosome 14 DNA sequence BAC R-109N23 of library RPc1-11
 ACCESSION AL442663
 VERSION AL442663.3 GI:15384728
 KEYWORDS HTG: HTGS_ACTIVEIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 165423)
 Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
 Brothier, P., Catolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
 Levy, M., Eckenberg, R., Bruls, T., deBardins, V., Crnaud, C.,
 Gysay, G., Saurin, W. and Weissenbach, J.
 Sequencing of the human chromosome 14
 Unpublished
 2 (bases 1 to 165423)
 Genoscope.
 Direct Submission
 Submitted (22-AUG-2001) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
 Web : www.genoscope.cns.fr)
 On Aug 31, 2001 this sequence version replaced gi:11191214.

 Center: Genoscope / Centre National de Sequencage
 Center code: GS
 Web site: http://www.genoscope.cns.fr/
 Contact: seque@genoscope.cns.fr

 The following BAC sequence is oriented from the T7 to the SP6 end.

 Summary Statistics
 Assembly program: Phrap; version 2.0
 Quality coverage: 10.76x in Q20 bases; sum-of-contigs

 Overall quality chart :
 Range : bases
 0 - 9 : 77
 10 - 19 : 35
 20 - 29 : 186
 30 - 39 : 184
 40 - 49 : 543
 50 - 59 : 2999
 60 - 69 : 5734
 70 - 79 : 7215
 80 - 89 : 18522
 90 - 99 : 45746
 84182

 Percentage of bases with a quality value >= 40 : 99 %.
 Location/Qualifiers
 1..165423
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone="R-109N23"
 /clone_1fb="RPc1-11"
 1347..1555
 /note="matching EMBL:U06118
 Rdbd:RH3746
 dbSTS:STS19225
 identified using the e-PCR software (G. Schuler)"
 4250..4436
 /note="matching EMBL:AA234560
 Rdbd:RH91914
 dbSTS:STS64918
 identified using the e-PCR software (G. Schuler)"
 134484..134584
 /note="matching EMBL:Z44639
 Rdbd:RH26227
 dbSTS:STS19666

Identified using the e-PCR software (G. Schuler)"

```

QY 4605 CTCAAAGCAACAAATATGTATTATCTAAACGTAAGCTATTAATCAGATATATGA .4664
DB 71002 GAGAAATCTCTGTCTCAAAAACAAACAAACAAACAAACAAACAAACAAACAAAT 71061
QY 4665 TAGCATTTTAAATTGAAAAAGCATTAAT-----GATTACATGATTTGTAAAA 4711
DB 71062 CACCTGACCAACCAACCAATCAGAACTCATCTGGGCTGAAGAGCCCATCAGAACTGAGCTA 71121
QY 4712 TATCAAAATACATGAAATCTGTGTCTTAATATATGCTAGCAACAAAGCCACATTTGTTT 4771
DB 71122 TATGACCAATCAGAACTAGCTAGTTGACGCCCTTCATTTGCATTAATATGAGCCGATG 71181
QY 4772 TTACTAGGCGACCAAG--GTACTTTAAAAAAGTTAGGCCACCAAG-----GGCTC 4824
DB 71182 GACTTGGGCGAGAACTTGTCTATAAAAAGCCGAACGTGTCAGCGGACGGGATGGCTC 71241
QY 4825 ACACCTGTATATCCAGCACTTTGGAGGCCAA-----GGCAGAGATCACTGAGGCC 4878
DB 71242 ACATCTGTATATCCAGCACTTTGGAGGCCAAAGGGGTGGGGGACATCTGAGGTC 71301
QY 4879 AGGAGTTAGACCTGAGCAACATAGGAGATCTGTCTCTATATAAAATTTAA 4938
DB 71302 AGGAGTTGAGACCACTGAGCAACATAGGAGATCTGTCTCTATATAAAATTTAA 4938
QY 4939 AAATGCGTAGGCCCTTTGGCTTACACCCGTAATCCAGCACTTTGGAGGCCGAGGC 4996
DB 71359 AAATTCAGAGACCTGTGTGTGTCACCTGTAGTCCAGATATCTGAGGAGCTGAGGC 71416

```

```

RESULT 13
AC006345 125508 bp DNA linear PRI 09-MAY-2001
LOCUS Homo sapiens PAC clone RP4-791N16 from 14q24.3, complete sequence.
DEFINITION AC006345
AC006345.4 GI:11120928
VERSION HTG.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125508)
AUTHORS Sulston, J.E. and Waterston, R.
JOURNAL Toward a complete human genome sequence
MEDLINE Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 99063792
9847074
REFERENCE 2 (bases 1 to 125508)
AUTHORS Scott, K., Layman, D., Hawkins, M. and Gregory, S.
JOURNAL The sequence of Homo sapiens PAC clone RP4-791N16
REFERENCE 3 (bases 1 to 125508)
AUTHORS Waterston, R.H.
JOURNAL Direct Submission
TITLE Submitted (11-JAN-1999) Genome Sequencing Center, Washington
UNIVERSITY University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 125508)
AUTHORS Waterston, R.
JOURNAL Direct Submission
TITLE Submitted (08-NOV-2000) Department of Genetics, Washington
UNIVERSITY University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 125508)
AUTHORS Waterston, R.
JOURNAL Direct Submission
TITLE Submitted (09-MAY-2001) Department of Genetics, Washington
UNIVERSITY University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Nov 8, 2000 this sequence version replaced g1:7630766.
-----
Center: Washington University Genome Sequencing Center
Genome Center
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.edu
-----
Summary Statistics

```

Center project name: H_DJ0791N16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
This clone from chromosome 14 was provided by Dr. Pieter de Jong, Roswell Park Cancer Institute, Human Genetics Department, Elm and Carlton Streets, Buffalo NY 14263-0001 USA.

SOURCE INFORMATION:
This clone was derived from human PAC library RPI-4, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://pacpac.med.buffalo.edu) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (http://www.genomesystems.com) or Research Genetics, Inc. (http://www.resgen.com); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RPI-14613, 200 base pair overlap. Actual start of this clone is at base position 1 of RP4-791N16; actual end is at base position 125312 of RP4-791N16.

The sequence in the following regions was derived from PCR product of RP4-791N16 PAC DNA: base positions 45895 to 46245, 64121 to 64136, 72649 to 72822, 89272 to 89446, 96124 to 96418, and 103487 to 103499.

FEATURES
There are polymorphic base pair differences in the overlap between the clone RP4-791N16 and RPI-14613.

```

source 1..125508
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="14"
        /map="14q24.3"
        /clone="RP4-791N16"
        /clone_1lb="RPI-4"
        .377
        /rpt_family="Alu"
        .568
        /rpt_family="Alu"
        .916
        /rpt_family="Alu"
        .1145
        /rpt_family="L1"
        .1628
        /rpt_family="Alu"
        .2265
        /rpt_family="Alu"
        .2553
        /rpt_family="Alu"
        .2865
        /rpt_family="Alu"
        .3034
        /rpt_family="Alu"
        .3567
        /rpt_family="Alu"
        .4430
        /rpt_family="L2"

```


[illegible][illegible]

RPI-309K20 is from the library RPI-1 constructed by the group of
Pleier de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

FEATURES

VECTOR: pcypac2
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.

Location/Qualifiers

1..38599
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone_lib="RPI-1"
/clone_id="RPI-1"
696..743
/note="24 copies 2 mer ta 93% conserved"
3725..4309
/note="match: GSS: Em: A0533081"
/complement(5117..12828)
/gene="d3309K20.1"
complement(join(<5117..5237,5324..5495,5664..5772,
6114..6475,7083..7269,7357..7540,7658..7817,8074..8302,
8561..>8692))
/gene="d3309K20.1"
/note="match: ESTs: Em: A1741534 Em: A1206080 Em: A1189057
Em: R74490 Em: R74601
match: proteins: Tr: O75923 SW: P08121 SW: P43027 Tr: O77086
SW: P29400 Tr: O9Y408 Tr: O92744 Tr: O95836 Tr: O929H0"
/codon_start=1
/evidence=not_experimental
/product="d3309K20.1.1 (novel protein similar to
dysterlin, isoform 1)"
/protein_id="CAB87608.1"
/db_xref="GI:7573588"
/translation="GVTVLEAKLVGININPVANVQGGRRVTRTQGTSCPFNE
VLEFPHDRLELDLELITVSGVVISLDRGDEKAGLTPSPKAFHSTLTFP
AARIGTFRMDLGIILDDPDQGYQKRWPLHDPDRAGTGLIKVLSVARAGDLPEP
MLPAPGHSIDIEKMLLPRGVAPERPARLRYRLRAGDLPALRLGLGSLYRALHD
ORLVEPYRVAFLEGOEGEYSAAEAAPENMOISFVLEPPLTRSLRLQDLDLPL
IDALATVHPDLRRIISHPRAGFPNTPGPAWPIYGGSPGAGIRDSIOLINGVCG
IMPRGLIIVASQVLEGRAPPEPPQAOGSLSTRTRKKKARDQTPKAVPHLD
ASPGAGPEIPRAMEVEBELPLPENVLAPCEDFLRGVLEATIMIDPTVASQPLSF
EISIRAGLELEQDRGSRAGGTEGAEAQPLGARPEEREKEELGTAAORPEPM
DGSGRYFCLPCHCKPCMHVMSCEMDHTWRLOSSNCVRVAERL"
complement(join(<6253..6475,7083..7269,7357..8302,
8561..8692,8810..8870,9348..9518,9855..9896,10081..10172,
12790..>12828))
/gene="d3309K20.1"
/product="d3309K20.1.2 (novel protein similar to
dysterlin, isoform 2)"
/note="novel protein similar to dysterlin
match: cDNAs: Em: A0526266
match: ESTs: Em: A0451439 Em: A1189057"
/evidence=not_experimental
complement(6253)
/gene="d3309K20.1"
6349..7896
/note="CpG island"
/evidence=not_experimental
complement(join(8051..8302,8561..8692,8810..8870,
9348..9367))
/gene="d3309K20.1"
/codon_start=1
/evidence=not_experimental
/product="d3309K20.1.2 (novel protein similar to
dysterlin, isoform 2)"
/protein_id="CAC33500.1"
/db_xref="GI:13235330"

repeat_region
/note="21 copies 2 mer ca 76% conserved"
10381..10408
/note="14 copies 2 mer tt 89% conserved"
complement(12015..12450)
/gene="d3309K20.1"
/note="match: GSS: Em: A0409243"
complement(19148..19617)
/note="match: GSS: Em: A0625097"
complement(19179..19617)
/note="match: GSS: Em: A0625350"
19633..20081
/note="match: GSS: Em: B86354
match: STS: Em: B86354"
20992..21045
/note="27 copies 2 mer aa 77% conserved"
21308..22898
/note="CpG island"
/evidence=not_experimental
21443..26600
/gene="SPAG4"
join(21443..21858,22687..22791,23071..23137,23330..23391,
23638..23681,24001..24026,24163..24271,24474..24549,
24746..24861,25130..25297,26235..26324,26427..26600)
/gene="SPAG4"
/product="d3309K20.2 (Sperm associated antigen 4
(acrosomal protein ACR55))"
/note="acrosomal protein ACR55 (similar to rat sperm
antigen 4 (SPAG4))"
match: cDNAs: Em: AF043345
match: ESTs: Em: A1635301 Em: AA742708 Em: A1168223
Em: A1537754 Em: A1347363 Em: A1183593 Em: A812953
Em: A1623233 Em: A0021925 Em: A0039264 Em: A873389"
/evidence=not_experimental
join(21555..21858,22687..22791,23071..23137,23330..23391,
23638..23681,24001..24026,24163..24271,24474..24549,
24746..24861,25130..25297,26235..26324,26427..26573)
/gene="SPAG4"
/note="match: proteins: Tr: O94901 Tr: O75156"
/codon_start=1
/evidence=not_experimental
/product="d3309K20.2 (Sperm associated antigen 4
(acrosomal protein ACR55))"
/protein_id="CAB87609.1"
/db_xref="GI:7573589"
/translation="MRRSRPGSASSSRKHTPNFSENSMSITSEDSKGLRSPGP
GPEGRARGPGSGEPALSGVGGTMMGSSQOKPAPRSHNNOTACGATVNGASE
PNCSPVSEERDILPTDLROMPPRPYKKSLLFQSLVLSLGGVIVSMYVE
VCSIRLFTVSLSLFLSFWGLGLVYVPSLENPKEMLTSEIHERVRSQOQLOD
LQAEIDKLKRETVSRANSEVAKLVFORLNDPVRKPYALSSVGASIDLOKTSDD
VADRNTAYFWNRSEFNWYARPTVILEPHEPNCAPFEDDOQVQLDGRVOLSIT
TLOHPPVSEHNGSAPRDEFAVLEQVDEFEVSGKTFPEVSEKIOTFHLNDP
PAAPKVKIOLISNMGHPRTCYLRVARGVRSSEGAEGSAQCPH"
complement(22956..23315)
/note="match: GSS: Em: A0284440"
complement(22969..23403)
/note="match: GSS: Em: A0466236"
/gene="SPAG4"
/note="match: STS: Em: 264454"
24633..25310
/gene="SPAG4"
/note="CpG island"
/evidence=not_experimental
24658..24701
/note="22 copies 2 mer cc 75% conserved"
26569..26574
/gene="SPAG4"
26596

[illegible][illegible]

Db	18832	ACTTGATCCCAAGAAAGGAGAGGTTGGCGGTAGGCCAAGATTGTCCACTGCATCCAGCCT	18773
QY	4591	GAGTGCAGAGCCTATCTC-----AAAGCAACAAATAATGTTATCTTAACGGTA	4641
Db	18772	GAGCCACAGAGTGAAGACTCTGTCTCCAAAAAAGAAAAAAGAAAAAACAACAC	18713
QY	4642	AGGTATTAATCAGAAATATATGATAGCATTTTAATTGAAAAAGCAATTAAATGATG	4701
Db	18712	ACATTTTGAAGAACAAATGCAGTATGTCATGACACAAAGTGCACAGCTAAGGCCAAACAGTCC	18653
QY	4702	GATTGTAAATATCAATATACATATAAATTTCTTGTTCTTAATATATCTATGACAACAGCA	4761
Db	18652	AGGTAGAGAAATACACTGATCATCATTTGAATATCCCAAATCTCGCCCTCGAGTGCATGC	18593
QY	4762	CATTGGTTTTCTAGAGGCCACCAAGATCTTTAAAAAAGTTAGGCGCCAGCCACAGGGG	4821
Db	18592	CTCTTAATCTCCACTCTCACAACTTCCCTTAGAATGTGAGGAGCGTGGGGCCGGCATATGACGG	18533
QY	4822	CTCAGACCTGTATATCCAGCACTTTGGGAGGCCAAGCAGAGATCACTTTAGCCCAAG	4881
Db	18532	CTCAGCGCTTACTCCAGCACTGTGTGAGACCCAGCA-GTGGATCACTTAAAGGTCAAG	18474
QY	4882	AGTTTAGAGCCTGAGCAACATAGGAGAGATCCTGATCTTCTCTATAAAAATTTAAAAA	4941
Db	18473	AGTTCAAGACCAACCTGGCCAAATATGTGTAATCTCATG--TCTACTAAAAATTAACAAA	18416
QY	4942	TTGGCTAGAGCCCTTTGGCTTTACACCCGTAATCCACACTTTTGGAGGCCAGGGCG	4998
Db	18415	TTTAGCTGGGTGGTGGCGCGTGTGCCCTGTATGTTTCAACTTCTTGGGAGGCTGTAGCGAG	18359

RESULT 15	LOCUS	DEFINITION	LOCUS	DEFINITION
HSJD7689N3/C	121502 bp	DNA	linear	PRI 31-JAN-2002
HSJD7689N3	Human DNA sequence from clone RP4-669N3 on chromosome Xp11.1-1.4			
	Contains the 5' end of the RP2 gene for retinitis pigmentosa 2 (X-linked recessive), the 5' end of the SLC9A6 gene for solute carrier family 9 (sodium/hydrogen exchanger) isoform 6 and a Cpg island, complete sequence.			
ACCESSION	AL050307			
VERSION	AL050307.13	GI:6562569		
KEYWORDS	HNG; Cpg Island; retinitis pigmentosa; RP2; SLC9A6; solute carrier family 9.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 121502)			
AUTHORS	Lawlor,S.			
TITLE	Direct Submission			
JOURNAL	Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,			

COMMENT

numQuery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Dec 12, 1999 this sequence version replaced 91:6016948.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence is
the entire insert of clone R4-869N. This sequence was generated
from part of bacterial clone contigs of human chromosome X,
constructed by the Sanger Centre Chromosome X Mapping Group.
Further information can be found at
<http://www.sanger.ac.uk/HGP/chrX>
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP-669N3 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/dacpac/home.htm>
VECTOR: pCYPAC2.

FEATURES	source	location/Qualifiers
		1..121502
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="X"
		/map="p11.1-11.4"
		/clone="RP4-689N3"
		/clone.lib="RPC1-4"
		990..1287
repeat_region		/note="MER82 repeat: matches 351..652 of consensus"
		1734..1933
repeat_region		/note="MER82 repeat: matches 145..354 of consensus"
		3656..3767
repeat_region		/note="L1M4A4 repeat: matches 5428..5539 of consensus"
		4209..4303
repeat_region		/note="L1M4A4 repeat: matches 5539..6165 of consensus"
		11604..11715
repeat_region		/note="56 copies 2 mer aa 60% conserved"
gene		complement(11740..34902)
		/gene="RP2"
mRNA		complement(join(<11740..11854,17701..18366,34640..34902))

CDS

/gene="RP2"
/product="dj689N3.1 (retinitis pigmentosa 2 (X-linked recessive))"
/note="match: CDNAS: Em:AJ007590
match: ESTs: Em:AL121304"
/evidence-not_experimental
complement(301n<11740. .11854,11701. .18366,34640. .34741))"
/gene="RP2"
/note="match: proteins: Tr:075695"
/codon_start=1
/evidence-not_experimental
/product="dj689N3.1 (retinitis pigmentosa 2 (X-linked recessive))"
/protein_id="CAB82030.1"
/db_xref="GI:7320873"
/db_xref="SPTREMBL:Q9NU67"
/translation="MGCFPSRRKADKREPERNEERPKOYSDMDREKVDPKDFMSG
LKDEVLGRPGVGAOFLDIODCENCNIYFDHATVITDCTGIIFLDPSGSPF
NAGRCRCITACQOFRVDRCKRLEFLCAGAOPIRESSSNKKGCCPMQYBELAFQK
DRCSTIFENNTSNIHDPFVSGELNMSLIPEDAVYQYVYPTTEELKARVSTAEAR
STVPSKQQRKRSDESCVLVLFAGDYITIANARKLIDEMVGKGFVLQYKVSMAKD
AORVREARPDFPLPLN"
23885. .24286
repeat_region
/note="201 copies 2 mer at 69% conserved"
24288. .24329
repeat_region
/note="21 copies 2 mer at 76% conserved"
26821. .27654
repeat_region
/note="L1P42 repeat: matches 5316. .6146 of consensus"
29352. .29356
repeat_region
/note="L1M44 repeat: matches 6094. .6294 of consensus"
29558. .29626
repeat_region
/note="L1M44 repeat: matches 5984. .6052 of consensus"
complement(32542. .32643)
/gene="RP2"
/note="match: GSS: Em:AQ381314"
32641. .33111
misc_feature
/note="match: GSS: Em:AQ453772"
34202. .35135
misc_feature
/note="Cpg island"
/evidence-not_experimental
35440. .35479
repeat_region
/note="20 copies 2 mer ct 77% conserved"
38874. .39187
misc_feature
/note="match: GSS: Em:AQ565300"

Db	73804	CGTGAAGTAGAGGATTCAAGACACCGCTGGCCAACTGSGCAAAACCCCTGTCTTACTTAA	737453
QY	3589	GCTGGGCTCCTGACACGAGCTCCTGGGCACTGGAGCTGAGGCTGCGCGCTGGGGGCGTCA	3648
Db	73744	AAATCAAAAATTAGCCAGCGGTGTGTGTGTGCTGTAAATCCAGCTACTTGGAGGCT	736855
QY	3649	GACTAAAGAGTCAACACTGAGAGGCTCTCAAGCAGGCCAGAGGTTTAAAGCACCTT	3708
Db	73684	GAG-----	73682
QY	3709	TCCTCCAAACCCAGAGAGACCCCTGGAGCCAGGCTTGTCTGGGCCACTCTACTGGCC	3768
Db	73681	-----GCAAGAGAAATCACTGTGAACCTGGGAGGCGGACCGCAGTACCCGAGAT	736522
QY	3769	TGTTTTACTGAAATCCACACAGACTATAGGCCACATAGTACATTAAAAAGAGAGAGA	3828
Db	73631	TGCACCACTGACACTGACGCTGTGATAGAAATAGAGCTGTGTCTCAAAAAA	73572
QY	3829	GAGAGAGAGAGAGAGAGATGAGATCTCACTGTGTGTCCAGGCTGTCTCGAATCTT	3888
Db	73571	AAAGTACAGAGATAG-----GAGTCTTCTTTTGTGGCCAGGTTGTCTTGAACCTCT	73518
QY	3889	AGGCTCAAGCAATCCCTGCTGCTTATGAGCTCCCAAGGGGCGGATTTACAGGTGAGGTA	3948
Db	73517	GGGCTCAAGAGATTCCTTCTTCCCTTCCAAAGTGTGAGATTACAGGTGAGGCA	73458
QY	3949	CTGCACTTGACACCA-----CATGATCTTTT	3978
Db	73457	CCACATCTGGCCGTGAATAATTTTTTACAGGGGCGCTATATTATTATTATTATTATTA	73398
QY	3979	TTTTTTTTTTTTTTTTTGGAGAGGTTTCACTCCATCAACCCAGGCTGGAGTGCAGTGG	4038
Db	73397	TCATTATTAATATTATTGACACAGGATCTTGTCTGTGCCAGGCTGGAGTGCAGT-G	73339
QY	4039	GGGCAATCTTGGCTACTGTAACTGTGCTGCCAGGCTGGAAGGATTCCTGCTTAG	4098
Db	73338	GGGTGATCTCAGCTCAATTGCAACCTGCTGCTGTGTGACAGAAATTAATCTGTCTCAG	73279
QY	4099	CCTCCGAGTACTGTGAATTATAGGCACACACACAGCGCTGGCTAATTTTTTTTTTT	4158
Db	73278	CCTCCGAGTACTGTGGACTACAGGGCGACGCTGCCACACCACTAATTTTTTTGT---	73222
QY	4159	TCGTATTTTTAGTAGACAGGGTTTCAATCATGTTGGCCAGCTGTGCTTGAACCCCTG	4218
Db	73221	-----ATTTATGTAAGATGGGGTTTCACTAATGTTATCCAGGCTGTCTGAACTCCTG	73168
QY	4219	ACCTCAAGTATCCACCCACCTGCGGCTCCCAAGTCTGGATTAAGGTGACAGCAC	4278
Db	73167	ACCTCAAGGAAATCCACCGGCTCAAGCTCCCAAGTCTGGATTAAGGTGTAAGCAC	73108
QY	4279	CATGCACAGCCCAATGTAATTTTTTAAATTTTAAATTTAAATGTTATCTAA	4338
Db	73107	TGTGTGGGGCCATGAGCATGTATTTTAAAGGCAAGTGA--ACAGTTGGGAAAAAGA	73051
QY	4339	GGCCAGTAGCATGACTGCGCTGTGTAATCCACGACTTTGAGGGCCAAAGTCCGGGA	4398
Db	73050	GGCTGGGTCAATGTGCTATGCTCTGTAATCCACGCTTTTGGGAGGCGCAAGAGAGA	72991
QY	4399	TCACTTGAAGCTGGGAGT-----CAGGTGGGCAACATAGTAGAGCCCGCTCTAC	4452
Db	72990	TCCCTTGAAGCCCAAGATTGAGACACCAAGCTGGGCAATAGGAGAGCCCGCTCTAC	72931
QY	4453	AAAAATTTAAAAATTTAGTGGGAGTGGGCAATTTGCCGTGTGCCACTACTTGGGA	4512
Db	72930	CAAAAATTAACAAAGTTAGCTGGGTGTGGGTGTGTGCTGTGTCCAACTCTTTGGGA	72871
QY	4513	AGCTAGAGTGGGAGTGGCTGAAGCTGTGAGGTGAGGCTGCAATGACTATGATCAC	4572
Db	72870	GGCCAAAGTGAAGAGACCTGAGTCCAGGAATTTGAGGACACAGTGAAGCTGTGTCTAT	72811
QY	4573	ACCAATGCACTCAGGCTAGTGAAGCAGGCTATCTCAAAAGCAAAATAATGTTTATC	4632
Db	72810	GCACTGCACTCAGGCTGGGTGACAGAGTGAAGCTCTGTCTCAAAAAA-----	72759

OY	4653	TAAACGGAGAGTTCATATACAGAAATTCATGATAGCATTTTAAATGAAAAAGCATTAAT	4692
Db	72758	-----MAAAAAGAAAAAAATTCATTAATAAATCTCTATTCCCTGAAAGCCTTAGAGCT	72707
OY	4693	GATTACATGGATTGTAAATAATCAAAATACATGAATAATCTTGCTTTTAATTA---TGC	4748
Db	72706	TTACAGAGACHTGAATCATCATCAAAATTAATGAGCGAGAGTGCAACTTAACAGCATTTGCC	72647
OY	4749	TAGCAACAAGGCACATTGGCTTTTCTACTAGGCGACCAAGTACTTTAAAAAAAGTTAGGG	4808
Db	72646	TTGAAAGGTGGGAAAGGAGAGAGTCTTAAATAATGCTACCTTTGAAAAATGTCATCATATA	72587
OY	4809	CCACACACAGGGGCTACACCTGTAATCCAGACACTTTGGAGAGCCCAAGCAGAGAGATC	4868
Db	72586	GCCAGCATGATGTCTCATGCTCTGTAAATCCAGCATTTGGAGAGCCCAAGCATGACGATT	72527
OY	4869	ACTTGAGCCCAAGAGATTAGGA----CCTGAGCCAAATAGGGAGA-TCCGTGATCTTGCT	4923
Db	72526	ATTTGAGGTGAGGAGGTTTGAGACCAAGCCTGGCCCAACATGTTGAAATCCATCTTTAATTA	72467
OY	4924	CTATAAAAAATTAAAAAATTTGGCTAGAGCCCTTTGGCTTAACCCGTAATCCAGCATTT	4983
Db	72466	AAATACAAACAACAAAAAATTTGGCTGTGTGTGGTGCACACACTGTAATCCAGCTACTC	72407
OY	4984	GGGAGGCCGAGGC 4996	
Db	72406	AGGAGGCTGGGCC 72394	

Search completed: June 8, 2003, 18:10:24
Job time : 12664.1 secs

DB 1275 GAGGACCTGCTGGCGAA 1292

RESULT 13

US-09-986-480-116
; Sequence 116, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P5500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 116
; LENGTH: 2076
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-986-480-116

Query Match 2.1%; Score 47; DB 9; Length 2076;
Best Local Similarity 81.5%; Pred. No. 0.0059;

Matches 53; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2165 CCCAACTAATGAGTTTAAATACAAATCTGTTAAAAA 2224
DB 1959 CCCAACTTGTCTGTATTAAGTCACTCTTTAAAAA 2018
QY 2225 AAAA 2229
DB 2019 AAAA 2023

RESULT 14

US-09-349-015-25
; Sequence 25, Application US/09349015A
; Patent No. US20020015950A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen Anne
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael
; TITLE OF INVENTION: ATHEROSCLEROSIS-ASSOCIATED GENES
; FILE REFERENCE: PB-0013 US
; CURRENT APPLICATION NUMBER: US/09/349,015A
; CURRENT FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 846
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-349-015-25

Query Match 2.1%; Score 46.6; DB 10; Length 846;
Best Local Similarity 66.3%; Pred. No. 0.0042;
Matches 67; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2129 TATCTGAATCCAGACACACCTTAGTATCTGCCAAATGAGTTTAAAT 2188
DB 511 TTGTGGCTCAAGTAGCTGATTTTAAATTTGTGAAATATAAATTTTCTTAT 570
QY 2189 ACAATCTCTGTTAAAAA 2229
DB 571 AGAAAAA 611

RESULT 15

US-10-156-239-7
; Sequence 7, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. US20030036074A1 Nucleic Acid Sequences Encoding Human Tr
; TITLE OF INVENTION: Atpase Molecule, A Human Ubiquitin Hydrolyase-Like Molecule, A
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 8195
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)...(7442)
US-10-156-239-7

Query Match 2.1%; Score 46.6; DB 9; Length 8195;
Best Local Similarity 68.8%; Pred. No. 0.019;
Matches 64; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 2137 TCAGACAGCCACCTTAGTATCTGCCAAATGAGTTTAAATACAAATAC 2196
DB 8074 TCCTGGGAGCCGCGCTGTACATAGGCGACAGATGTTGTTTAAATTAACAA 8133
QY 2197 TCGTTAAAAA 2229
DB 8134 TGTCAAAAAA 8166

Search completed: June 9, 2003, 05:31:17
Job time : 313.723 secs

XX Petrukhin K, Caskey CT, Metzker M, Madellius C;
XX WPI: 1999-540560/45.
XX P-PSDB; AAY29953.
XX Human and mouse polynucleotides encoding CGICE polypeptides
XX Claim 2; Fig 2; 67pp; English.
XX The present sequence represents the human CGICE cDNA sequence, which
XX when mutated is responsible for Best's macular dystrophy (BMD).
XX polynucleotides encoding CGICE are useful for diagnosing whether a
XX patient carries a mutation in the CGICE gene. Normal and mutated
XX CGICE proteins are useful for identifying activators and/or inhibitors
XX of these proteins, in order to treat BMD. The CGICE gene offers a
XX simpler and cheaper method of diagnosing BMD without the need for the
XX presence of the patient. The gene may also be useful to discovering
XX the genetic cause of age-related macular dystrophy.
SQ Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 other;
Query Match 100.0%; Score 2229; DB 20; Length 2229;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGAGTCCACCAAGCCTAGTCCGACAGCTTCTGTGGATCATCGAACCCACTGGA 60
DB 1 CAGGAGTCCACCAAGCCTAGTCCGACAGCTTCTGTGGATCATCGAACCCACTGGA 60
QY 61 CCCCACGACCCCAAGCCCACTGCTGACAGCCACTGCTGCTGCTGCTGCTGCTGCTG 120
DB 61 CCCCACGACCCCAAGCCCACTGCTGACAGCCACTGCTGCTGCTGCTGCTGCTGCTG 120
QY 121 CAAAGCAAGTGGCTATGCTCCGCTAGGCTCTCTCCGCTGCTGCTGCTGCTGCTG 180
DB 121 CAAAGCAAGTGGCTATGCTCCGCTAGGCTCTCTCCGCTGCTGCTGCTGCTGCTG 180
QY 181 GCAGCATATACAGTGTCTATATGCGAGTCTTATCTCTGCTGCTGCTGCTGCTGCT 240
DB 181 GCAGCATATACAGTGTCTATATGCGAGTCTTATCTCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGCTATTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 TCCGCTATTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 CTCTGATTTGGACAGCTACATCATCTATCCCATTTCTTCTGCTGCTGCTGCTGCT 360
DB 301 CTCTGATTTGGACAGCTACATCATCTATCCCATTTCTTCTGCTGCTGCTGCTGCT 360
QY 361 TGACGCTGTGCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 TGACGCTGTGCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 TCATAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TCATAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GCAGCCTATCCGCTACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GCAGCCTATCCGCTACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CAGTGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CAGTGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 CAGAACACAGAGTGGAGAACTGAGCTACACAAACATGTTGGGCTGCTGCTGCTG 660
DB 601 CAGAACACAGAGTGGAGAACTGAGCTACACAAACATGTTGGGCTGCTGCTGCTGCTG 660
QY 661 TGTGTTTCCCAACCTGTCAATGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 TGTGTTTCCCAACCTGTCAATGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 TGTCTCAGAGCTCTGATGAGATGAACACCTTGGCTACTCAGTGTGACACCTGTATG 780
DB 721 TGTCTCAGAGCTCTGATGAGATGAACACCTTGGCTACTCAGTGTGACACCTGTATG 780
QY 781 CCTACGACTGATATGATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 CCTACGACTGATATGATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 GCTTCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 841 GCTTCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 GGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GCTGCTGAGTGGCAGAGCAGCTCATCAACCCCTTGGAGAGATGATGATGATTTTG 1020
DB 961 GCTGCTGAGTGGCAGAGCAGCTCATCAACCCCTTGGAGAGATGATGATGATTTTG 1020
QY 1021 AGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 AGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 ACCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 ACCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 CCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 CCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 TCAGCTGAAACAAAGAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 TCAGCTGAAACAAAGAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 ACCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 ACCTGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 CAACTCAAGACCAACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1321 CAACTCAAGACCAACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1381 CCAAAACCAAGGACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1440
DB 1381 CCAAAACCAAGGACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1440
QY 1441 ACCTTAAGCTGTGACGCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 ACCTTAAGCTGTGACGCTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 GTGCCCCAGACGCCCCCTGACGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1501 GTGCCCCAGACGCCCCCTGACGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1561 CGTCAAGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1561 CGTCAAGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 GTTCTGGGCTCAAGAAATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1621 GTTCTGGGCTCAAGAAATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY 1681 ACCCAAGATATCTCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
DB 1681 ACCCAAGATATCTCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
QY 1741 AGATCCCGGAGAAATCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
DB 1741 AGATCCCGGAGAAATCACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800

QY 1801 CACTCAAGATCATGATGATCTTATTGGCCCTTGGAAAAAGGGATGAAGACACTTCT 1860
 DB 1801 CACTCAAGATCATGATGATCTTATTGGCCCTTGGAAAAAGGGATGAAGACACTTCT 1860
 QY 1861 AACCTCTCTCTATATGGGGATGCTTGCCGCCAGGATGCTTACCTGTGTACACCGCA 1920
 DB 1861 AACCTCTCTCTATATGGGGATGCTTGCCGCCAGGATGCTTACCTGTGTACACCGCA 1920
 QY 1921 GGACACTGATCCAGTACAGCCATACAGCTGTCCACAGTGAAGAGTGTCTACACAG 1980
 DB 1921 GGACACTGATCCAGTACAGCCATACAGCTGTCCACAGTGAAGAGTGTCTACACAG 1980
 QY 1981 CCTGAATCAAAATGGTATGCTTAATAGATATAAAATCCAGACTCTTCCAGCTTAATGCC 2040
 DB 1981 CCTGAATCAAAATGGTATGCTTAATAGATATAAAATCCAGACTCTTCCAGCTTAATGCC 2040
 QY 2041 TTTTATTCATTAATAACTGTGCAAGCTAGACTGAACCTTGGAAACATTTAACTCAGACTC 2100
 DB 2041 TTTTATTCATTAATAACTGTGCAAGCTAGACTGAACCTTGGAAACATTTAACTCAGACTC 2100
 QY 2101 TGGATTCAGAGTCGGGAAACCTTATGCTATCTGAATCCAAAGACAGCCACCTTAGTAT 2160
 DB 2101 TGGATTCAGAGTCGGGAAACCTTATGCTATCTGAATCCAAAGACAGCCACCTTAGTAT 2160
 QY 2161 ACTGCCCAACTATGATGTTTAAATAATACAAATCTCTTAAAAAAGGGGGGGGGGGGG 2220
 DB 2161 ACTGCCCAACTATGATGTTTAAATAATACAAATCTCTTAAAAAAGGGGGGGGGGGGG 2220
 QY 2221 AAAAAAAAA 2229
 DB 2221 AAAAAAAAA 2229

RESULT 2
 AA221228
 ID AA221228 standard; cDNA; 2429 bp.
 XX
 AC AA221228;
 XX
 DT 22-NOV-1999 (first entry)
 XX
 DE Human CG1CE long form cDNA sequence.
 XX
 KM CG1CE; Best's macular dystrophy; mutation; diagnosis; detection;
 XX BMD; age-related macular dystrophy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 105..1412
 FT /product- "CG1CE short form protein"
 FT
 FT
 PN MO9943695-A1.
 XX
 PD 02-SEP-1999.
 XX
 PF 22-FEB-1999; 99MO-US03790.
 XX
 PR 25-FEB-1998; 98US-0075941.
 PR 18-DEC-1998; 98US-0112926.
 XX
 PA (MERI) MERCK & CO INC.
 PA (UYUP-) UNIV DPPSALA.
 XX
 PI Petukhin K, Caskey CT, Metzker M, Wadelius C;
 XX
 DR MPI, 1999-540560/45.
 DR P-PSDB; AAY29954.
 XX
 PT Human and mouse polynucleotides encoding CG1CE polypeptides
 XX
 PS Claim 2; Fig 4; 67pp; English.

XX The present sequence represents the human CG1CE cDNA sequence, which
 CC when mutated is responsible for Best's macular dystrophy (BMD).
 CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
 CC patient carries a mutation in the CG1CE gene. Normal and mutated
 CC CG1CE proteins are useful for identifying activators and/or inhibitors
 CC of these proteins, in order to treat BMD. The CG1CE gene offers a
 CC simpler and cheaper method of diagnosing BMD without the need for the
 CC presence of the patient. The gene may also be useful to discovering
 CC the genetic cause of age-related macular dystrophy.
 XX
 SQ Sequence 2429 bp; 614 A; 694 C; 613 G; 508 T; 0 other;
 Query Match 90.3%; Score 2013; DB 20; Length 2429;
 Best Local Similarity 91.6%; Pred. No. 0;
 Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;
 QY 1 CAGGGAGTCCACACAGCTAGTGCAGACCTTCTGTGGATCATCGGACCCACTGGAA 60
 DB 1 CAGGGAGTCCACACAGCTAGTGCAGACCTTCTGTGGATCATCGGACCCACTGGAA 60
 QY 61 CCCACCTGACCCCAAGCCGCTGCTGAGGCCGCTGCGCATGACATCACTTACA 120
 DB 61 CCCACCTGACCCCAAGCCGCTGCTGAGGCCGCTGCGCATGACATCACTTACA 120
 QY 121 CAAGCCAAAGTGGCTAATGCCCGCTTAGGCTCTTCTCCGCTGCTGCTGCTGCGGG 180
 DB 121 CAAGCCAAAGTGGCTAATGCCCGCTTAGGCTCTTCTCCGCTGCTGCTGCTGCGGG 180
 QY 181 GCAGCATCTACAAAGCTCTATATGCGAGTCTTAACTTCTCTGCTGCTACATCA 240
 DB 181 GCAGCATCTACAAAGCTCTATATGCGAGTCTTAACTTCTCTGCTGCTACATCA 240
 QY 241 TCCGCTTATTTATAGCTGCGCCCTCACGGAAGAACAACAGCTGATTTAGAACTGA 300
 DB 241 TCCGCTTATTTATAGCTGCGCCCTCACGGAAGAACAACAGCTGATTTAGAACTGA 300
 QY 301 CTCTGTATTCGACAGCTATACATCCAGCTATCCCATTTCTTCTGCTGCTGCTACG 360
 DB 301 CTCTGTATTCGACAGCTATACATCCAGCTATCCCATTTCTTCTGCTGCTGCTACG 360
 QY 361 TGACGCTGTGCTGACCCGCTGCTGGAACAAGTACGAAACCTGCTGCTGCTGCTGCTG 420
 DB 361 TGACGCTGTGCTGACCCGCTGCTGGAACAAGTACGAAACCTGCTGCTGCTGCTGCTG 420
 QY 421 TCATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 421 TCATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GCAAGCTCATCCGCTACGCAACCTGCGCAACGCTCTATCTGCGAGCTCAGACCG 540
 DB 481 GCAAGCTCATCCGCTACGCAACCTGCGCAACGCTCTATCTGCGAGCTCAGACCG 540
 QY 541 CAGTCTACAAAGCGCTTCCCAAGCGCCACAGACCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 541 CAGTCTACAAAGCGCTTCCCAAGCGCCACAGACCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 CAGAACCAAGAGCTTGGAGAACTGAGCTTACCAACCAACATGTTGCTGCTGCTGCTG 660
 DB 601 CAGAACCAAGAGCTTGGAGAACTGAGCTTACCAACCAACATGTTGCTGCTGCTGCTG 660
 QY 661 TGTGTTTGGCAACCTGTCAATGAAGGCTGTGAGGCTGCAATCCGGGACCTATCC 720
 DB 661 TGTGTTTGGCAACCTGTCAATGAAGGCTGTGAGGCTGCAATCCGGGACCTATCC 720
 QY 721 TGTCTCAGAGCTGTCTGAAGAGATGAACACTTGTCTACTCACTGTGAGACCTGTATG 780
 DB 721 TGTCTCAGAGCTGTCTGAAGAGATGAACACTTGTCTACTCACTGTGAGACCTGTATG 780
 QY 781 CTTACGACTGATAGTATGATATCCCACTGCTGTATACAGAGTGTGCTGCTGCTGCTG 840
 DB 781 CTTACGACTGATAGTATGATATCCCACTGCTGTATACAGAGTGTGCTGCTGCTGCTG 840

```

OY 841 GCTTCTCCGATGCTGATGTTGGGCGGAGTTTCTGAACCCAGCCAGGAGGCTACCCCTG 900
    |||
Db 841 GCTTCTCCGATGCTGATGTTGGGCGGAGTTTCTGAACCCAGCCAGGAGGCTACCCCTG 900
OY 901 GCCATGAGTGAAGCTGTTGTCGCCGCTTTCAGGTTCTGCAGATTCTTCTATGTTG 960
    |||
Db 901 GCCATGAGTGAAGCTGTTGTCGCCGCTTTCAGGTTCTGCAGATTCTTCTATGTTG 960
OY 961 GCTGGGCTGAA----- 970
    |||
Db 961 GCTGGGCTGAAAGTGGGCGCTCTCCAGGCGCTGCTGGGCTGAGAGCATGGCCAGAGGGGTC 1020
OY 971 ----- 970
    |||
Db 1021 ATGGCCAGCAGCTGCTTGAACGAGATGTCAGTGTCTAGGAAGAAAGTCTCAGGGGTAG 1080
OY 971 ----- 970
    |||
Db 1081 AAGCAGCCAGGCGTGTGGCGCAGACACTGTAAATCCAGCTACTCGGAGGCTGAGGCGAG 1140
OY 971 ----- 970
    |||
Db 1141 GAGAAATCGCTTGAACCCGGAGGCGGAGTTGTGTGGCAGAGCAGCTCATCAACCCCTT 1200
OY 998 TGGAGAGGATGATGATGATTTTGGAGACCACTGGATTGTGCAGAGAAATTTGACGTGTC 1057
    |||
Db 1201 TGGAGAGGATGATGATGATTTTGGAGACCACTGGATTGTGCAGAGAAATTTGACGTGTC 1260
OY 1058 CCGTGTGGCTGTGGATGATGATGACACAGAACCTGCTCGATGAGGCGGAGCATGTACTG 1117
    |||
Db 1261 CCGTGTGGCTGTGGATGATGATGACACAGAACCTGCTCGATGAGGCGGAGCATGTACTG 1320
OY 1118 GAATTAAGCCCGAGCAGACAGCCCGCTACAGCTGCTCGCCAGTTCGCTGAGAGCTC 1177
    |||
Db 1321 GAATTAAGCCCGAGCAGACAGCCCGCTACAGCTGCTCGCCAGTTCGCTGAGAGCTC 1380
OY 1178 CTTTATGGGCTCCACCTTCAACATCAGCTGACCAAGAGAGAGATGAGTTCAGCCCAA 1237
    |||
Db 1381 CTTTATGGGCTCCACCTTCAACATCAGCTGACCAAGAGAGAGATGAGTTCAGCCCAA 1440
OY 1238 TCGAGAGGAGAGAGAGATGCTCAGCGTGGCATATTGGCCGCTTCTAGGCGTGCAGTC 1297
    |||
Db 1441 TCGAGAGGAGAGAGAGATGCTCAGCGTGGCATATTGGCCGCTTCTAGGCGTGCAGTC 1500
OY 1298 CCAATGATCAGCAATCTCCCAAGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGA 1357
    |||
Db 1501 CCAATGATCAGCAATCTCCCAAGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGA 1560
OY 1358 ATCCCTTCTCCAGAGAGGCGCTGCCCCAAAACACAGAGGAGCAGCAAGAACTTTAGGGG 1417
    |||
Db 1561 ATCCCTTCTCCAGAGAGGCGCTGCCCCAAAACACAGAGGAGCAGCAAGAACTTTAGGGG 1620
OY 1418 CCAAGGAAGACACAAAGGCTGGAAGCTTAAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACT 1477
    |||
Db 1621 CCAAGGAAGACACAAAGGCTGGAAGCTTAAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACT 1680
OY 1478 GTATCAGAGGCGCAGGCTACTACATAGTGGCCCAAGAGCGCCCTCAGGCCCATCCCATGT 1537
    |||
Db 1681 GTATCAGAGGCGCAGGCTACTACATAGTGGCCCAAGAGCGCCCTCAGGCCCATCCCATGT 1740
OY 1538 CTTTCCCTTGAAGACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGAGCATAGACACCAA 1597
    |||
Db 1741 CTTTCCCTTGAAGACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGAGCATAGACACCAA 1800
OY 1598 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAAAGAAAAGTTTGAATGCTCTCAGA 1657
    |||
Db 1801 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAAAGAAAAGTTTGAATGCTCTCAGA 1860
OY 1658 GAGCGATGGGGCGCTTGTATGAGGACCCAGAAAGTATCTCAATGAGAGGAGAAAGTGTGGA 1717
    |||
Db 1861 GAGCGATGGGGCGCTTGTATGAGGACCCAGAAAGTATCTCAATGAGAGGAGAAAGTGTGGA 1920
OY 1718 GTTAACTGACGATATGCGCAGAGATCCCGAAAATCACCCTCAAGAAAGCTTTGAGACA 1777

```

```

Db 1921 GTTAACTGACGATATGCGCAGAGATCCCGAAAATCACCCTCAAGAAAGCTTTGAGACA 1980
    |||
OY 1778 ATCAACCAACACATATACACTATACCTAAAGATCACATGATCTCTTATTTGGGCTTGGGA 1837
    |||
Db 1981 ATCAACCAACCAACATATACACTATACCTAAAGATCACATGATCTCTTATTTGGGCTTGGGA 2040
    |||
OY 1938 AAACAGGATGATGACATCTTCTAAGCTGCTTCCATATGGGATGCTTGGCAGCAGGT 1897
    |||
Db 2041 AAACAGGATGATGACATCTTCTAAGCTGCTTCCATATGGGATGCTTGGCAGCAGGT 2100
    |||
OY 1898 CCTCAGCTGTGTGACACAGCAGACAGACACTGATCCAGTACAGCCATACAGCTGTCCACA 1957
    |||
Db 2101 CCTCAGCTGTGTGACACAGCAGACAGACACTGATCCAGTACAGCCATACAGCTGTCCACA 2160
OY 1958 CTGAAGAACGTGCTCTTCAACAAAGCCTGAATGAATGTTAGCTTAAATGATTAANAATGCC 2017
    |||
Db 2161 CTGAAGAACGTGCTCTTCAACAAAGCCTGAATGAATGTTAGCTTAAATGATTAANAATGCC 2220
    |||
OY 2018 AGACTACTTCAAGCTTAACTGATGCTTAACTGATTAATAAAGCTGAAAGCTAGACTGACCA 2077
    |||
Db 2221 AGACTACTTCAAGCTTAACTGATGCTTAACTGATTAATAAAGCTGAAAGCTAGACTGACCA 2280
OY 2078 TTGGAACATTTAACTGAGCTGTGATTCAGAGTGGGAAACCTTAACTTATCTGAAT 2137
    |||
Db 2281 TTGGAACATTTAACTGAGCTGTGATTCAGAGTGGGAAACCTTAACTTATCTGAAT 2340
OY 2138 CCAAGACAGCCACACACTTATGATGATGCTGCCCCAAACTATGAGTAAATTAATCAAAATNT 2197
    |||
Db 2341 CCAAGACAGCCACACACTTATGATGATGCTGCCCCAAACTATGAGTAAATTAATCAAAATNT 2400
OY 2198 CGTTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 2226
    |||
Db 2401 CGTTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 2429
    |||

RESULT 3
AAV9722
ID AAV9722 standard; cDNA: 1263 BP.
XX
XX AAV9722;
XX
XX 26-APR-1999 (first entry)
XX
XX Human adult retina secreted protein bx112_15 cDNA.
XX
XX Secreted protein; human; retina; bx112_15; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 100..885
XX FT /*tag= a
XX
XX W0856909-A2.
XX
XX 17-DEC-1998.
XX
XX 08-JUN-1998; 98WO-US11822.
XX
XX 05-JUN-1998; 98US-0092722.
XX
XX 11-JUN-1997; 97US-0873218.
XX
XX (GENE ) GENETICS INST INC.
XX
XX Agostino MJ, Fechtel K, Howes SH, Jacobs K, LaVallie ER;
XX McCoy JM, Racie LA, Spaulding V, Treacy M;
XX
XX WPI: 1999-080899/07.
XX P-PSDB; AAV95345.
XX
XX New polynucleotides encoding secreted human proteins - derived from
XX human foetal brain, adult testes, foetal kidney, adult thyroid or

```

adult retina cDNA libraries

Claim 12: Page 71-72: 113pp: English.

This is the nucleotide sequence of cDNA clone bk112-15, which includes an open reading frame for a 261-amino acid polypeptide (see AA95345). The clone was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins, or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoding protein. Database searches indicate some sequence similarity to known sequences. The invention provides cDNA clones (see AA9721-33) from human adult thyroid, adult retina, adult testis, foetal kidney and foetal brain that encode novel secreted proteins (see AA95344-53). Each clone is individually available from deposit clone ATCC 98451 (see also AA9724-43). The isolated polynucleotides (PNS) and proteins are predicted to have activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional, cytokine, cell proliferation or differentiation, immune stimulating (e.g. as vaccines) or immune suppressing, haematopoiesis regulating, tissue growth, thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour invasion suppressor, and tumour inhibition activities. The PNS are also stated to be useful for gene therapy.

Sequence 1263 BP: 397 A; 342 C; 272 G; 252 T; 0 other:

Query Match 56.0%; Score 1247.2; DB 20; Length 1263;

Best Local Similarity 99.8%; Pred. No. 2e-281;

Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

978 GAGAGGTCATCAACCCCTTGGAGAGATATGATTTTGGAGACCACTGATGTC 1037
 1 GAGAGGTCATCAACCCCTTGGAGAGATATGATTTTGGAGACCACTGATGTC 60
 1038 GACAGAAATTTGAGAGTGTCTTGGCTGTGATGATGACACAGGAGCTGCTGG 1097
 61 GACAGAAATTTGAGAGTGTCTTGGCTGTGATGATGACACAGGAGCTGCTGG 120
 1098 ATGAGAGCCGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1157
 121 ATGAGAGCCGACATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 1158 GCCAGTTCGTCGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGAAACAA 1217
 181 GCCAGTTCGTCGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGAAACAA 240
 1218 GAGATGAGTTCAGAGCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
 241 GAGATGAGTTCAGAGCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 1278 CGCTTCCTAGAGCTGAGTCCATGATCAACATCTCCAGGGCAAACTCAAGAGCAA 1337
 301 CGCTTCCTAGAGCTGAGTCCATGATCAACATCTCCAGGGCAAACTCAAGAGCAA 360
 1338 CTACGTGGCCCAAGAGGAAATCCCTTCTCCAGAGAGGCTGCCCCAAGCAAGGCA 1397
 361 CTACGTGGCCCAAGAGGAAATCCCTTCTCCAGAGAGGCTGCCCCAAGCAAGGCA 420
 1398 GCCAAACAGAGCTTGGGGCCAGAGCAACAAAGGCTGGAAAGCTTAAAGCTGTGAG 1457
 421 GCCAAACAGAGCTTGGGGCCAGAGCAACAAAGGCTGGAAAGCTTAAAGCTGTGAG 480
 1458 GCTTCACAGTGTGGCCACATGATCAGAGAGGCTTACATCAGTGTGAGAGAGAG 1517
 481 GCTTCACAGTGTGGCCACATGATCAGAGAGGCTTACATCAGTGTGAGAGAGAG 540
 1518 CTCAGAGCCCACTCCATGTTCTCCCTAGAAACATCAGAGGCTCAAGAGCTTCAAGT 1577
 541 CTCAGAGCCCACTCCATGTTCTCCCTAGAAACATCAGAGGCTCAAGAGCTTCAAGT 600

QY 1578 GTCACAGCATAGACACCAAGCAAAAGCTTAAAGACTGTGAGTCTGGGCCCAAGAA 1637
 DB 601 GTACAGGCAATACACCAAGCAAAAGCTTAAAGACTGTGAGTCTGGGCCCAAGAA 660
 QY 1638 AGTTTGAATTTGCTCTCAGAGAGCGATGGGCTTGTAGAGCACCAGAAAGTATCA 1697
 DB 661 AGTTTGAATTTGCTCTCAGAGAGCGATGGGCTTGTAGAGCACCAGAAAGTATCA 720
 QY 1698 GTGAGAGAGAAACCTGTGAGTTAAACCTGACGATATGCGAGATGCCGAAATATC 1757
 DB 721 GTGAGAGAGAAACCTGTGAGTTAAACCTGACGATATGCGAGATGCCGAAATATC 780
 QY 1758 CTCAAAGAACCTTTGGAGAACATCACCAACCAATCAGCTACCTCAAGAAATCATG 1817
 DB 781 CTCAAAGAACCTTTGGAGAACATCACCAACCAATCAGCTACCTCAAGAAATCATG 840
 QY 1818 GATCCTTATTTGGGCTTGGAAAAACAGGATGAAGACATTTCTTACCTGCTCTAATG 1877
 DB 841 GATCCTTATTTGGGCTTGGAAAAACAGGATGAAGACATTTCTTACCTGCTCTAATG 900
 QY 1878 GGATGCTTGGCCAGGAGGCTCTCACTGCTGTGTACACGACGACATGATCAATG 1937
 DB 901 GGATGCTTGGCCAGGAGGCTCTCACTGCTGTGTACACGACGACATGATCAATG 960
 QY 1938 CAGCCATACAGCTGTCCACACTGAAGAACGTGTCTTACCAACAGCCGTAATCAATG 1997
 DB 961 CAGCCATACAGCTGTCCACACTGAAGAACGTGTCTTACCAACAGCCGTAATCAATG 1020
 QY 1998 GCTTAATGATTAATAATCCAGACTTACCTTACCTTAAATGCTTATTAATCAAT 2057
 DB 1021 GCTTAATGATTAATAATCCAGACTTACCTTACCTTAAATGCTTATTAATCAAT 1080
 QY 2058 GTGAAGCTTAAGCTGACACTGAAGAACGTGTCTTACCAACAGCCGTAATCAATG 2117
 DB 1081 GTGAAGCTTAAGCTGACACTGAAGAACGTGTCTTACCAACAGCCGTAATCAATG 1140
 QY 2118 ACCCTTATTTATCTGAATCCAGACGACACTTAAATGATGATGATGATGATG 2177
 DB 1141 ACCCTTATTTATCTGAATCCAGACGACACTTAAATGATGATGATGATGATG 1200
 QY 2178 GTTAAATTAATCAATATCTGTTAAAAAATTTTTTTTTTTTTTTTTTTTTT 2229
 DB 1201 GTTAAATTAATCAATATCTGTTAAAAAATTTTTTTTTTTTTTTTTTTTTT 1252

RESULT 4
 AA221229
 ID AA221229 standard; cDNA: 1916 BP.
 XX
 AC AA221229;
 XX
 AC 22-NOV-1999 (first entry)
 XX
 DE Mouse CG1CE cDNA sequence.
 XX
 KW CG1CE; Best's macular dystrophy; mutation; diagnosis; detection;
 XX BMD; age-related macular dystrophy; ss.
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 11..1666
 FT /tag- a
 FT /product- "CG1CE protein"
 XX
 PD W09943695-A1.
 XX
 PD 02-SEP-1999.
 XX
 PF 22-FEB-1999; 99MO-US03790.
 XX
 PR 25-FEB-1998; 9805-0075941.

PR 18-DEC-1998; 98US-0112926.
 XX (MERI) MERCK & CO INC.
 PA (UYUP-) UNIV UPPSALA.
 XX Petrukhin K, Caskey CT, Metzker M, Madellius C;
 XX WPI: 1999-540560/45.
 DR P-PSDB; AAY29955.
 XX Human and mouse polynucleotides encoding CGICE polypeptides
 PS Claim 2; Fig 8; 67pp; English.
 XX The present sequence represents the mouse CGICE cDNA sequence, which
 CC when mutated is responsible for Best's macular dystrophy (BMD).
 CC Polyucleotides encoding CGICE are useful for diagnosing whether a
 CC patient carries a mutation in the CGICE gene. Normal and mutated
 CC CGICE proteins are useful for identifying activators and/or inhibitors
 CC of these proteins, in order to treat BMD. The CGICE gene offers a
 CC simpler and cheaper method of diagnosing BMD without the need for the
 CC presence of the patient. The gene may also be useful to discovering
 CC the genetic cause of age-related macular dystrophy.
 XX Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 other:
 SQ
 Query Match 39.1%; Score 871.2; DB 20; Length 1916;
 Best Local Similarity 77.1%; Pred. No. 1.7e-193;
 Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;

QY 756 CGTACTAGTGTGACACACCTGTATGCTACAGACTGGATTAGTATCCCACTGGTGTATACA 815
 DB 662 CGTACTAGTGTGACACACCTGTATGCTACAGACTGGATTAGTATCCCACTGGTGTATACA 721
 QY 816 CAGGTGGTACGTGTGGGTGATACAGCTTCTCTGACTGTCTGTAGTGGTGGGAGGAGTT 875
 DB 722 CAGGTGGTACGTGTGGGTGATACAGCTTCTCTGACTGTCTGTAGTGGTGGGAGGAGTT 781
 QY 876 CTGAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGACCTGATGTCGCTTTCAGC 935
 DB 782 CTGAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGACCTGATGTCGCTTTCAGC 841
 QY 936 TTCTGCACTTCTTCTATGTGTGGCTGGCTGGAAGTGGGACAGACAGCTATCAACCC 995
 DB 842 ATCCGCAATTCCTTCTATGTGTGGCTGGCTGGAAGTGGGACAGACAGCTATCAACCC 901
 QY 996 TTGGAGAGGATGATGATGATTTTGAACAACCTGATGTCGACAGAAATTTGACAGTG 1055
 DB 902 TTGGAGAGGATGATGATGATTTTGAACAACCTGATGTCGACAGAAATTTGACAGTG 961
 QY 1056 TCCTGTTGGCTGTGATGATGATGACACAGACCTGCTGGATGAGCCGACATGAC 1115
 DB 962 TCCTGTTGGCTGTGATGATGATGACACAGACCTGCTGGATGAGCCGACATGAC 1021
 QY 1116 TGGATTAAGCCGACACACAGCCCTTACACAGCTGCTTCCGACATTCGTCGAGCC 1175
 DB 1022 TGGATTAAGCCGACACACAGCCCTTACACAGCTGCTTCCGACATTCGTCGAGCC 1081
 QY 1176 TCCTTAAGGCTGACCTTCAACATCAAGCTGACAAAGAGAGTGGATTCAGCC 1235
 DB 1082 TCCTTAAGGCTGACCTTCAACATCAAGCTGACAAAGAGAGTGGATTCAGCC 1141
 QY 1236 AATCAGAGAGAGAGAGAGAGTCAAGCTG-----GCATCATTTGGCCGCTTC 1283
 DB 1142 AAGAGAGAGAGAGAGAGAGTCAAGAGAGTGTATAGACAGACCTAGCTGCTTC 1201
 QY 1284 CTAGGCTGACAGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343
 DB 1202 TTAGAGCTGACACCAAAATCAACATCTCTCTTGAAGAGCTTGAAGAGCTATG 1261
 QY 1344 TGGCCCAAG 1403
 DB 1262 TGTCTTGA-----ACCCCTCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1315
 QY 1404 CAGAACGTTAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1463
 DB 1316 CAGAACG-----ATGCTGGAATTTAAGAGTGTGAGCTGCTTC 1354
 QY 1464 AAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523
 DB 1355 AATGTGTTCCAGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1414
 QY 1524 CCCACTCC 1531
 DB 1415 CACCTAC 1422

RESULT 5
 AA221226
 ID AA221226 standard; DNA; 16125 BP.
 XX AA221226;
 AC 22-NOV-1999 (first entry)
 XX
 DE Human CGICE genomic DNA sequence.
 XX CGICE: Best's macular dystrophy; mutation; diagnosis; detection;
 KW BMD; age-related macular dystrophy; ss.
 XX Homo sapiens.
 OS

Key	Location/Qualifiers
FT Intron	1..1513
FT Intron	/tag- a
FT Intron	/number- 1
FT Intron	1514..1570
FT Intron	/tag- b
FT Intron	/number- 1
FT Intron	1571..1578
FT Intron	/tag- c
FT Intron	/number- 2
FT Intron	1579..1586
FT Intron	/tag- d
FT Intron	/number- 2
FT Intron	1587..1594
FT Intron	/tag- e
FT Intron	/product= "CGICE protein"
FT Intron	/note= "contains introns"
FT Intron	1595..1602
FT Intron	/tag- f
FT Intron	/number- 3
FT Intron	1603..1610
FT Intron	/tag- g
FT Intron	/number- 3
FT Intron	1611..1618
FT Intron	/tag- h
FT Intron	/number- 4
FT Intron	1619..1626
FT Intron	/tag- i
FT Intron	/number- 4
FT Intron	1627..1634
FT Intron	/tag- j
FT Intron	/number- 5
FT Intron	1635..1642
FT Intron	/tag- k
FT Intron	/number- 5
FT Intron	1643..1650
FT Intron	/tag- l
FT Intron	/number- 6
FT Intron	1651..1658
FT Intron	/tag- m
FT Intron	/number- 6
FT Intron	1659..1666
FT Intron	/tag- n
FT Intron	/number- 7
FT Intron	1667..1674
FT Intron	/tag- o
FT Intron	/number- 7
FT Intron	1675..1682
FT Intron	/tag- p
FT Intron	/number- 8
FT Intron	1683..1690
FT Intron	/tag- q
FT Intron	/number- 8
FT Intron	1691..1698
FT Intron	/tag- r
FT Intron	/number- 9
FT Intron	1699..1706
FT Intron	/tag- s
FT Intron	/number- 9
FT Intron	1707..1714
FT Intron	/tag- t
FT Intron	/number- 10
FT Intron	1715..1722
FT Intron	/tag- u
FT Intron	/number- 10
FT Intron	1723..1730
FT Intron	/tag- v
FT Intron	/number- 11
FT Intron	1731..1738
FT Intron	/tag- w
FT Intron	/number- 11

W09943695-A1.

XX	02-SEP-1999.
PD	22-FEB-1999; 99WO-US03790.
XX	
PF	25-FEB-1998; 98US-0075941.
XX	18-DEC-1998; 98US-0112926.
PR	(MERI) MERCK & CO INC.
XX	(UYUP-) UNIV UPSALA.
PA	Petukhin K, Caskey CT, Metzger M, Wadellius C;
XX	WPI; 1999-540560/45.
DR	P-PSDB; AAY29953.
XX	
PT	Human and mouse polynucleotides encoding CGICE polypeptides -
XX	Claim 2; Fig 1; 67pp; English.
PS	
XX	
CC	The present sequence represents the human CGICE gene, which when
CC	mutated is responsible for Best's macular dystrophy (BMD).
CC	Polynucleotides encoding CGICE are useful for diagnosing whether a
CC	patient carries a mutation in the CGICE gene. Normal and mutated
CC	CGICE proteins are useful for identifying activators and/or inhibitors
CC	of these proteins, in order to treat BMD. The CGICE gene offers a
CC	simpler and cheaper method of diagnosing BMD without the need for the
CC	presence of the patient. The gene may also be useful to discovering
CC	the genetic cause of age-related macular dystrophy.
XX	
SQ	Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 21 other;
XX	
Query Match	28.8%; Score 642.2; DB 20; Length 16125;
Best Local Similarity	99.5%; Pred. No. 1.3e-139;
Matches	644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1202 CAGCCTGAACAAGAGAGATGAGATTCCAGCCCAATCAGAGAGAGAGATGCTCA 1261
DB	13916 CAGCCTGAACAAGAGAGATGAGATTCCAGCCCAATCAGAGAGAGATGCTCA 13975
QY	1262 CGGTGGCATCAATGGCCCTTCCTAGGCTGAGTCCATGATCCATCCCTCCAGGGC 1321
DB	13976 CGGTGGCATCAATGGCCCTTCCTAGGCTGAGTCCATGATCCATCCCTCCAGGGC 14035
QY	1322 AAACCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCACAGAGGCTGCC 1381
DB	14036 AAACCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCACAGAGGCTGCC 14095
QY	1382 CAAAAACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1441
DB	14096 CAAAAACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 14155
QY	1442 GCTTAAGGCTGTGACAGCCTTCAAGTCTGCGCCCACTGATATAGAGGCCAGCTACTACAG 1501
DB	14156 GCTTAAGGCTGTGACAGCCTTCAAGTCTGCGCCCACTGATATAGAGGCCAGCTACTACAG 14215
QY	1502 TGCCCCACAGAGCGCCCTTCAGCCCACTCCATCTTCTCCCTAGAACATCAGCGCC 1561
DB	14216 TGCCCCACAGAGCGCCCTTCAGCCCACTCCATCTTCTCCCTAGAACATCAGCGCC 14275
QY	1562 GTCAAAGCTTCACAGTGTCAAGCATAGACACCAAGGACCAAGGACCAAGGACCAAGG 1621
DB	14276 GTCAAAGCTTCACAGTGTCAAGCATAGACACCAAGGACCAAGGACCAAGGACCAAGG 14335
QY	1622 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCTTGAATGAGCA 1681
DB	14336 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCTTGAATGAGCA 14395
QY	1682 CCCAGAGATATCTCAAGTGTGAGAGAAAGTGTGAGTTTAACCTGAGGATATGAGCA 1741
DB	14396 CCCAGAGATATCTCAAGTGTGAGAGAAAGTGTGAGTTTAACCTGAGGATATGAGCA 14455
QY	1742 GATCCCCGAAATACACTCAAGAACCTTTGAAACAATCAACCAACCAATACACTAC 1801

Db 14456 GATCCCGGAATATCACTCAAGAACCTTTGGAAACAATCACCACCAATACACTAC 14515
QY 1802 ACTCAAGATCAGATGATCTTATTTGGCCCTTGAAAAAGGATG 1848
Db 14516 ACTCAAGATCAGATGATCTTATTTGGCCCTTGAAAAAGGATCTG 14562
RESULT 6
ABAL4559
ID ABAL4559 standard; DNA; 16650 BP.
XX
AC ABAL4559;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6890.
XX
KM Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
KM immunosuppressive; anti-infective; anti-HIV; antibacterial; vulnary;
KM antiparkinsonian; antischistosomal; antileishmanial; cancer;
KM antitubercular; hepatotropic; cerebroprotective; anti-inflammatory;
KM antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KM neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN NC200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001KO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.

PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249268.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259676.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPT: 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system
cancers and metastases -
XX
PS Disclosure; SEQ ID NO 6888; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABAI1004-ABR21534) and proteins
CC (ABBI678-ABBI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;

Query Match	28.8%	Score 642.2	DB 22	Length 18530
Best Local Similarity	99.5%	Prod. No. 1.4e-139		
Matches 644	Conservative	0	Mismatches 3	Indels 0
				Gaps 0

QY	1202	CAGCCTCAACAAAGAGAGATGGAGTTCCAGGCCAATCAGAGAGACGAGGAGATGCTCA	1261
Db	11895	CAGCCTCAACAAAGAGAGATGGAGTTCCAGGCCAATCAGAGAGAGAGAGATGCTCA	11954
QY	1262	CGCTGGCATTTGGCCGCTTCTTAGGCCGCGAGTCCATGATCACCATCTCCAGGGC	1321
Db	11955	CGCTGGCATTTGGCCGCTTCTTAGGCCGCGAGTCCATGATCACCATCTCCAGGGC	12014
QY	1322	AAACTCAAGACCAAACTACTGTGGCCCAAGAGGAAATCCTTCTCAGAGGGCCCTGCC	1381
Db	12015	AAACTCAAGACCAAACTACTGTGGCCCAAGAGGAAATCCTTCTCAGAGGGCCCTGCC	12074
QY	1382	CAAAAAACAAAGGAGCCAAACAGAACGTTAGGGGCGAGGAAGAACAAAGGCCCTGGAA	1441
Db	12075	CAAAAAACAAAGGAGCCAAACAGAACGTTAGGGGCGAGGAAGAACAAAGGCCCTGGAA	12134
QY	1442	GCTTAAAGCTGTGGAGCCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGTACTACAG	1501
Db	12135	GCTTAAAGCTGTGGAGCCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGTACTACAG	12194
QY	1502	TGCCCCACAGAGCCCTCAGCCCACTCCCAATGTTCTTCCGCTTGAACCAATCAGCGCC	1561
Db	12195	TGCCCCACAGAGCCCTCAGCCCACTCCCAATGTTCTTCCGCTTGAACCAATCAGCGCC	12254
QY	1562	GTCAAAGCTTCACAGTGTACAGAGGCATAGACACAAAGACAAAGCTTAAACACTGTGAG	1621
Db	12255	GTCAAAGCTTCACAGTGTACAGAGGCATAGACACAAAGACAAAGCTTAAACACTGTGAG	12314
QY	1622	TTCCTGGGGCCAAAGAAAGTTTGAATGCTCTTCAGAGAGCGATGGGGCCTTGATGAGCA	1681
Db	12315	TTCCTGGGGCCAAAGAAAGTTTGAATGCTCTTCAGAGAGCGATGGGGCCTTGATGAGCA	12374
QY	1682	CCGAGAGATCTCAAGTGTAGAGAGGAAAACTGTGGATTAACTGAGCGATATGCGCAGA	1741
Db	12375	CCGAGAGATCTCAAGTGTAGAGAGGAAAACTGTGGATTAACTGAGCGATATGCGCAGA	12434
QY	1742	GATCCCCGAAAAATCACCTCAAGAACCTTTGGAAACAATCCCAACCAATATCACACTAC	1801
Db	12435	GATCCCCGAAAAATCACCTCAAGAACCTTTGGAAACAATCCCAACCAATATCACACTAC	12494
QY	1802	ACTCAAGATCACTGGATCCCTATTGGCCCTTGGAAAAACAGGATG 1848	
Db	12495	ACTCAAGATCACTGGATCCCTATTGGCCCTTGGAAAAACAGGATG 12541	

RESULT 8
ABA14558

AC ABA145587

DT 23-JAN-2002 (first entry)
 22

Human nervous system related polynucleotide SEQ ID NO 6889,

antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;

XX antihypertensive; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticoagulant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds
 XX
 OS Homo sapiens.
 NN WO200159063-A2.

Homo sapiens.

WO200159063-A2.

16-AUG-2001

PF 17-JAN-2001; 2001WO-US013334.

PR	04-FEB-2000	2000US-0180628
PR	04-FEB-2000	2000US-0180628
PR	24-FEB-2000	2000US-0184664
PR	02-MAR-2000	2000US-0186500
PR	16-MAR-2000	2000US-0189574
PR	17-MAR-2000	2000US-0190076
PR	18-MAR-2000	2000US-0191823
PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214866
PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	23-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227182
PR	08-SEP-2000	2000US-0228124
PR	01-SEP-2000	2000US-0228287
PR	01-SEP-2000	2000US-0228287
PR	01-SEP-2000	2000US-0228344
PR	01-SEP-2000	2000US-0228344
PR	05-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	06-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0229517
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231410
PR	08-SEP-2000	2000US-0231411
PR	12-SEP-2000	2000US-0232081
PR	14-SEP-2000	2000US-0232197
PR	14-SEP-2000	2000US-0232338
PR	14-SEP-2000	2000US-0232339
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	14-SEP-2000	2000US-0233065
PR	14-SEP-2000	2000US-0233066
PR	14-SEP-2000	2000US-0233066
PR	14-SEP-2000	2000US-0233067
PR	14-SEP-2000	2000US-0233067
PR	14-SEP-2000	2000US-0233068
PR	14-SEP-2000	2000US-0233068
PR	14-SEP-2000	2000US-0233069
PR	14-SEP-2000	2000US-0233069
PR	14-SEP-2000	2000US-0233070
PR	14-SEP-2000	2000US-0233070
PR	14-SEP-2000	2000US-0233071
PR	14-SEP-2000	2000US-0233071
PR	14-SEP-2000	2000US-0233072
PR	14-SEP-2000	2000US-0233072
PR	14-SEP-2000	2000US-0233073
PR	14-SEP-2000	2000US-0233073
PR	14-SEP-2000	2000US-0233074
PR	14-SEP-2000	2000US-0233074
PR	14-SEP-2000	2000US-0233075
PR	14-SEP-2000	2000US-0233075
PR	14-SEP-2000	2000US-0233076
PR	14-SEP-2000	2000US-0233076
PR	14-SEP-2000	2000US-0233077
PR	14-SEP-2000	2000US-0233077
PR	14-SEP-2000	2000US-0233078
PR	14-SEP-2000	2000US-0233078
PR	14-SEP-2000	2000US-0233079
PR	14-SEP-2000	2000US-0233079
PR	14-SEP-2000	2000US-0233080
PR	14-SEP-2000	2000US-0233080

QY 1682 CCCAGAGTATCTCAAGTGAAGAGGAAAACTGTGAGTTTAACTGACGATATGCCAGA 1741
 |||||
 Db 12381 CCCAGAGTATCTCAAGTGAAGAGGAAAACTGTGAGTTTAACTGACGATATGCCAGA 12440
 QY 1742 GATCCCGGAAATTCACCTTCAAAAGAACCTTTGGAACATCAACCAACATACACACTAC 1801
 |||||
 Db 12441 GATCCCGGAAATTCACCTTCAAAAGAACCTTTGGAACATCAACCAACATACACACTAC 12500
 |||||
 QY 1802 ACTCAAGATCAATGATGATCTTATTTGGGCTTTGGGAAACAGGATG 1848
 |||||
 Db 12501 ACTCAAGATCAATGATGATCTTATTTGGGCTTTGGGAAACAGGATG 12547
 |||||
 RESULT 9
 ABL89697 standard: cDNA, 1717 BP.
 ID ABL89697 standard: cDNA, 1717 BP.
 XX
 AC ABL89697;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 259.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antifibrotic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PE 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Blaise CE, Rosen CA;
 XX
 DR WPI; 2002-122018/16.
 XX
 DR P-PSDB; ABB89288.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 4; SEQ ID NO 259; 2081pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions, e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 9 other;

Query Match 26.4%; Score 587.8; DB 24; Length 1717;
 Best Local Similarity 98.8%; Pred. No. 2, 7e-127;
 Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 QY 1242 GAGGACGAGAGGAGATGCTCAGCGTGCATCTATTGGCCCTTCTAGCCCTGACATCCAT 1301
 |||||
 Db 1 GAGGACGAGAGGAGATGCTCAGCGTGCATCTATTGGCCCTTCTAGCCCTGACATCCAT 60
 |||||
 QY 1302 GATCACCATTCTCCCGAGGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAATCC 1361
 |||||
 Db 61 GATCACCATTCTCCCGAGGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAATCC 120
 |||||
 QY 1362 CTCTCCACAGAGGAGGCTGCCAAAACACACAGGACCCCAACAGAAAGCTTAGGGGCGAG 1421
 |||||
 Db 121 CTCTCCACAGAGGAGGCTGCCAAAACACACAGGACCCCAACAGAAAGCTTAGGGGCGAG 179
 |||||
 QY 1422 GAAGACACAGAGGCTGGAAGCTTAAAGCTGTGGAGCCCTTCAAGTCTGGCCCACTGTAT 1481
 |||||
 Db 180 GAAGACACAGAGGCTGGAAGCTTAAAGCTGTGGAGCCCTTCAAGTCTGGCCCACTGTAT 239
 |||||
 QY 1482 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGAGCCCTCAGCCCACTCCATGTTCTTC 1541
 |||||
 Db 240 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGAGCCCTCAGCCCACTCCATGTTCTTC 299
 |||||
 QY 1542 CCCCTAGAACATCAAGCGCCGCTCAAAAGCTTCAACAGTGTACAGAGCATAGACCAAAAGAC 1601
 |||||
 Db 300 CCCCTAGAACATCAAGCGCCGCTCAAAAGCTTCAACAGTGTACAGAGCATAGACCAAAAGAC 359
 |||||
 QY 1602 AAAAGCTTAAGACATGTGATGTTGGGGCCCAAGAAAAGTTTGAATGCTCTCAGAGAGC 1661
 |||||
 Db 360 AAAAGCTTAAGACATGTGATGTTGGGGCCCAAGAAAAGTTTGAATGCTCTCAGAGAGC 419
 |||||
 QY 1662 GATGGGCTTATGATGAGAGCAGCCAGAGATGTCAAGTGAAGAGGAAAGTGGGAGTT 1721
 |||||
 Db 420 GATGGGCTTATGATGAGAGCAGCCAGAGATGTCAAGTGAAGAGGAAAGTGGGAGTT 479
 |||||
 QY 1722 AACCTGACGAGATATCCAGAGATGCCGAAAATCACTTCAAGAACCTTTGGAAACATCA 1781
 |||||
 Db 480 AACCTGACGAGATATCCAGAGATGCCGAAAATCACTTCAAGAACCTTTGGAAACATCA 539
 |||||
 QY 1782 CCACCAACATACACATCACTCAAAAGTCAATGATCTTATTTGGGCTTTGAAAAAC 1841
 |||||
 Db 540 CCACCAACATACACATCACTCAAAAGTCAATGATCTTATTTGGGCTTTGAAAAAC 599
 |||||
 QY 1842 AGGAGTG 1848
 |||||
 Db 600 AGGAGTG 606
 |||||
 RESULT 10
 ABA14556
 ID ABA14556 standard: DNA; 7108 BP.
 XX
 AC ABA14556;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 6887.
 XX
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
 KW antiparkinsonian; antidiabetic; antianemic; antithyroidic; cancer;
 KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antifibrotic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW neurotoxic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.
 PF
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-MAR-2000; 2000US-0198123.
 PR 19-MAR-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217467.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 XX

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251858.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02359678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-541565/60.
 DR
 XX

PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure: SEQ ID NO 6887; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AB11004-AB21533) and proteins
CC (AB14678-AB21800) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 other;

Query Match 21.4%; Score 476.4; DB 22; Length 7108;
Best Local Similarity 99.4%; Pred. No. 5.1e-101;
Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1202 CAGCCTGACAAAGAGAGATGAGTTCACGCCCAATAGAGAGCGAGAGATGCTCA 1261
DB 6609 CAGCCTGACAAAGAGAGATGAGTTCACGCCCAATAGAGAGCGAGAGATG-TC A 6667
QY 1262 CGCTGAGCATATGCGCGCTCTAGAGCCTGAGTCCATATGACATCCCTCCAGGCG 1321
DB 6668 CGCTGAGCATATGCGCGCTCTAGAGCCTGAGTCCATATGACATCCCTCCAGGCG 6727
QY 1322 AACTCAAGAGACCAACTACTGTGCGCCCAAGAGAGGAAATCCCTTCCAGAGGCGTCC 1381
DB 6728 AACTCAAGAGACCAACTACTGTGCGCCCAAGAGAGGAAATCCCTTCCAGAG- G 6786
QY 1382 CAAAAACCAAG 1441
DB 6787 CAAAAACCAAG 6846
QY 1442 GCTTAAAGCTGTGAGAGCCTTCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACTCAG 1501
DB 6847 GCTTAAAGCTGTGAGAGCCTTCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACTCAG 6906
QY 1502 TCGCCCAAG 1561
DB 6907 TCGCCCAAG 6966
QY 1562 GTCAAGAGCTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
DB 6967 GTCAAGAGCTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7026
QY 1622 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 7027 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 7086
QY 1682 CCCAGAGATATCTCAAGTGAAG 1703
DB 7087 CCCAGAGATATCTCAAGTGAAG 7108

RESULT 11
AB10793
ID AB10793 standard; cDNA; 2861 BP.
XX
AC AB10793;

XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmacological; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2001; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PERK) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB66690.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.
XX
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (AB16176-AB16351), expressed DNA
sequences (AB57737-AB72072),
CC (AB57737-AB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other;

Query Match 14.1%; Score 314.8; DB 23; Length 2861;
Best Local Similarity 57.4%; Pred. No. 2.1e-63;
Matches 656; Conservative 0; Mismatches 462; Indels 24; Gaps 4;

QY 103 CCATGACATCACTTACACAGCCAGAGG---CTAATGCCCGCTTAGGCTCTCTCC 159
DB 413 CAAATGACATCACTTACACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472
QY 160 GCGTGTGCTGTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 219
DB 473 AATGCTGTGCTGTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
QY 220 TCTGT 279
DB 533 TCTGT 592
QY 280 AGCTGATGTTTGAAGAACTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
DB 593 AAGAAACCTTTGAGGCGCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
QY 340 CCTGT 399
DB 653 CCTGT 712
QY 400 ACCTGCGT 459
DB 713 CCATTCCTGT 772

QY 460 AGCAGGCGGCTGCTGCGGCGACGCTCATCCGCTACGCCAACCTGGCACTGCTCA 519
DB 773 AGCGAGGACCATGATGAGCGCAATATGCGATGTGTGCTTGGCTGACGATGG 832
QY 520 TCCTGGCGACGCTCAGCAGCCGAGTCTACAGCGCTTCCCGAGCGCCAGCAGCTGGTGC 579
DB 833 TCCTGGCGAATGTTTGGCCGAGGGTGAAGAGCGTTTCCCGGCTTAATATGTGGTGG 892
QY 580 AACGAGCGTTTATGACTCCGCGAGACACAGCAGTTGAGAAAGTACG-----C 630
DB 893 AACGCGGCTGCTAAATGACAAATGAAAGACATCATGAGACATGAACAAAGCGCTTTC 952
QY 631 TACACAGCAATGTTCTGGTGGCGCCGTTGGTGTGTTTGGCAACCTTCATAGAGCGGT 690
DB 953 CCAGACCTTGAAGACACTGCTGCTGCCATGCTTGGGCTGCGAGTATATTAACCGAGGCCA 1012
QY 691 GCGTTGGAGTCAATCCGAGCCATCTCTGCTCCAGAGCTGCTGACAGAGATGACA 750
DB 1013 GAAGAGAGGTCCGATTCGTGATGATTTGCTGGAAGACCATCATGATGAGCTAATA 1072
QY 751 CCTGGGCTACTAGTGTGGACACTGATGCTTACGACCTGAGATTAGTATCCACTGGTGT 810
DB 1073 AGTTTGGTGTGAGTGTGAGTCTCTCATACGATACGATACGATAGTACTCTGCTGT 1132
QY 811 ATACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 870
DB 1133 ACACCCAGAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1192
QY 871 AGT-----TTTGAAACCCAGCAAGGCTTACCTGCGCATGAGCTGAGCTGCTTG 921
DB 1193 AATGAGCCGATGSCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1252
QY 922 TGGCCGCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 981
DB 1253 TTCTGTATTTACAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1312
QY 982 AGCTCATCAACCCCTTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041
DB 1313 CGCTGATTAATCATTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
QY 1042 GCAATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1101
DB 1373 GCAATCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1432
QY 1102 AGCGGACATGTACTGGAATAGCCGAGCC---ACAGCCCCCTTACAGCTGCTTCCG 1158
DB 1433 TAAAGATCAGTACTGGAGAGAGTGTTCCTCAAGAGCTGCTTACCAATAGCTGCCG 1492
QY 1159 CCCAGTTCGCTGAGCCTCTTATGAGGCTCAACCTTCAACATCAGCTGACAAAGAGG 1218
DB 1493 AACGATTCGCGGAGATCATCCAGAGCCGTCACCTGCCAAGATGAGGTGCCAAGATG 1552
QY 1219 AG 1220
DB 1553 CG 1554

RESULT 12
ABAI1094
ID ABAI1094 standard; cDNA; 305 BP.
XX
XX ABAI1094;
DT 23-JAN-2002 (first entry)
XX
XX Human nervous system related polynucleotide SEQ ID NO 101.
XX
XX Human; nootropic; neuroprotective; cytosolic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparthenonin; antistaking; antianaemic; antithyroidic; cancer;
KW antileumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;

KW antiparastlic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
OS
XX Homo sapiens.
FN WO200159063-A2.
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232387.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

Db 141 GATATGCCAGAGATCCCCGAAAAATCACCCTTGAACAATCACCACCAC 200
QY 1791 ATACACCTACCTCAAAAGATCAGATGCTTATTTGGCCTTGAACACAGGATG 1848
Db 201 ATACACCTACCTCAAAAGATCAGATGCTTATTTGGCCTTGAACACAGGCTG 258

Search completed: June 8, 2003, 11:23:47
Job time : 452.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:29:42 ; Search time 91.8557 Seconds
(Without alignments)
7441.920 Million cell updates/sec

Title: US-09-622-964-2
Perfect score: 2229
Sequence: 1 cagggaggtccaccagccta.....aaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/prodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/prodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/prodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/prodata/1/lna/PCBUS.COMB.seq: *
6: /cgn2_6/prodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.6	1.9	1212	4	US-09-182-145-34 Sequence 34, App1
2	41.6	1.9	1212	4	US-09-182-145-35 Sequence 35, App1
3	41.2	1.8	1161	1	US-08-086-439C-2 Sequence 2, App1
4	41.2	1.8	1161	1	US-08-434-877-2 Sequence 2, App1
5	41.2	1.8	1367	3	US-08-475-742-3 Sequence 3, App1
6	41.2	1.8	1370	1	US-08-056-051-1 Sequence 1, App1
7	41.2	1.8	1370	1	US-07-928-611-17 Sequence 17, App1
8	41.2	1.8	1370	2	US-08-487-811A-17 Sequence 17, App1
9	41.2	1.8	1370	4	US-09-060-694-17 Sequence 17, App1
10	41.2	1.8	1370	4	US-09-378-074-17 Sequence 17, App1
11	41.2	1.8	1370	5	US-09-378-074-17 Sequence 17, App1
12	41.2	1.8	1466	1	US-08-056-051-3 Sequence 3, App1
13	41.2	1.8	1466	1	US-07-928-611-19 Sequence 3, App1
14	41.2	1.8	1466	2	US-08-487-811A-19 Sequence 19, App1
15	41.2	1.8	1466	4	US-09-060-694-19 Sequence 19, App1
16	41.2	1.8	1466	4	US-09-378-074-19 Sequence 19, App1
17	41.2	1.8	1466	5	US-09-378-074-19 Sequence 19, App1
18	41.2	1.8	1507	4	US-09-453-323-1 Sequence 1, App1
19	41.2	1.8	1507	1	US-08-056-051-5 Sequence 5, App1
20	41.2	1.8	1610	1	US-07-928-611-21 Sequence 21, App1
21	41.2	1.8	1610	2	US-08-487-811A-21 Sequence 21, App1
22	41.2	1.8	1610	4	US-09-060-694-21 Sequence 21, App1
23	41.2	1.8	1610	4	US-09-378-074-21 Sequence 21, App1
24	41.2	1.8	1610	5	US-09-378-074-21 Sequence 21, App1
25	41.2	1.8	1700	3	US-08-897-340-4 Sequence 4, App1
26	41.2	1.8	1700	3	US-08-897-340-4 Sequence 4, App1
27	41.2	1.8	7218	1	US-08-232-463-14 Sequence 14, App1

28	40.8	1.8	1194	4	US-08-765-907A-9 Sequence 9, App1
29	40.8	1.8	1233	4	US-08-765-907A-12 Sequence 12, App1
30	40.8	1.8	1561	4	US-08-765-907A-11 Sequence 11, App1
31	40.8	1.8	3124	4	US-09-734-030-1 Sequence 1, App1
32	40.8	1.8	4496	4	US-08-765-907A-6 Sequence 6, App1
33	40.6	1.8	2634	4	US-09-463-238-3 Sequence 3, App1
34	40.6	1.8	3275	4	US-09-370-838-151 Sequence 151, App
35	40.4	1.8	1512	2	US-08-909-965C-8 Sequence 8, App1
36	40.4	1.8	1697	4	US-09-346-408-5 Sequence 5, App1
37	40.2	1.8	458	1	US-08-524-757-1 Sequence 262, App
38	40.2	1.8	593	4	US-09-385-982-262 Sequence 9, App1
39	40.2	1.8	976	2	US-08-504-459-9 Sequence 6, App1
40	40.2	1.8	1373	5	PCT-US92-02977-6 Sequence 3, App1
41	40.2	1.8	1373	5	PCT-US95-03032-3 Sequence 27, App1
42	39.8	1.8	969	2	US-08-365-486A-27 Sequence 27, App1
43	39.8	1.8	969	4	US-08-880-342-27 Sequence 2, App1
44	39.8	1.8	8051	2	US-08-576-626A-2 Sequence 674, App
45	39.6	1.8	724	4	US-08-998-416-674

ALIGNMENTS

```
RESULT 1
US-09-182-145-34
; Sequence 34, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gueney, Austin J.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Penicka, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 34
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-34

Query Match          1.9%: Score 41.6; DB 4; Length 1212;
Best Local Similarity 73.6%: Pred. No. 0.23;
Matches 53; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 2158 TATCTGCCCAACTAATGATTAAATACAAATACCTGTTAAAAA 2217
Db 1119 TATATATAAAAAATAGTAAATAAAAA 1178

OY 2218 AAAAAAAAAA 2229
Db 1179 AAAAAAAAAA 1190

RESULT 2
US-09-182-145-35/C
; Sequence 35, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
```

```

APPLICANT: Botstein, David A.
APPLICANT: Cohen, Robert
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Lawrence, David A.
APPLICANT: Levine, Arnold J.
APPLICANT: Pennica, Diane
APPLICANT: Roy, Margaret Ann
APPLICANT: Wood, William I.
TITLE OR INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: P117682
CURRENT APPLICATION NUMBER: US/09/182,145B
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: US 60/063,704
EARLIER FILING DATE: 1997-10-29
EARLIER APPLICATION NUMBER: US 60/073,612
EARLIER FILING DATE: 1998-02-04
EARLIER APPLICATION NUMBER: US 60/081,695
EARLIER FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 156
SEQ ID NO 35
LENGTH: 1212
TYPE: DNA
ORGANISM: Homo sapiens
US-09-182-145-35

```

```

Query Match          1.9%; Score 41.6; DB 4; Length 1212;
Best Local Similarity 73.6%; Pred. No. 0.23;
Matches 53; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      2158 TATAGTCCCAACTAATGACGTTTAATAATACAAATCTCGTTAAAAA 2217
       ||||| . ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db      94 TATATATAAAAAATTAGAGCTAAAAA 35
QY      2218 AAAAAAAAAA 2229
       |||||||||
Db      34 AAAAAAAAAA 23

RESULT 3
US-08-086-439C-2
Sequence 2, Application US/08086439C
Patent No. 5468615
GENERAL INFORMATION:
APPLICANT: Chio, Christopher L.
APPLICANT: Huff, Rita M.
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
TITLE OF INVENTION: Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSSEE: The Upjohn Company, Corp. Intellectual
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0,, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,439C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darinley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4700
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-385-5210

```

```

1 TELEFAX: 616-385-6697
2
3 TELEX: 224401
4
5 INFORMATION FOR SEQ ID NO: 2
6
7 SEQUENCE CHARACTERISTICS
8
9 LENGTH: 161 base pairs
10
11 TYPE: nucleic acid
12
13 STRANDEDNESS: single
14
15 TOPOLOGY: linear
16
17 MOLECULE TYPE: cDNA
18
19 US-08-086-439C-2

```

Query Match	1.88; Score 41.2; DB 1; Length 1161;
RefSeq: G141444	60.0%; Prod No 0.38;

Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 430 TGTGTCGGGCTTCGTCGAGCAGGACAGCAGCGCCGGCTGTCTCGCGGCGACGCTCA 489
Db 257 TGTGCGCCGCTTCTCTCTATCTCCAGGTCAGGGTGGCGCGCTGGCTGCTGACCCCCCCC 316

```

Oy      490 TCCGCTACGGCAACTGGGCAACGTGTCATCTCCGCGACGGTGACACC 539
          | | | | | | | | | | | | | | | | | | | | | |
Db      317 TGTGGACGCGCCATGCGCATGGACATGCTGTGACACGCGCTTCATC 366

```

RESULT 4
US-08-434-877-2
: Sequence 2, Application US/08434877

GENERAL INFORMATION:

APPLICANT: Huff, Rita M.

TITLE OF INVENTION: Receptors

CORRESPONDENCE ADDRESS:

ADDRESSEE: Property Law

CITY: Kalamazoo

COUNTRY: USA

COMPUTER READABLE FORM:

COMPUTER: Gateway 2000, P5-90

SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/434,877
FILING DATE: 1 July 1992

CLASSIFICATION: 435

NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33 673

REFERENCE/DOCKET NUMBER: 4700 DV1
 ATTORNEY INFORMATION:

TELEPHONE: 616-385-5210
CITY FAX: 616-385-6007

TELEX: 224401

SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid

TOPOLOGY: 1linear

MS-08-434-877-2

Query Match	1.88;	Score 41.2;	DB 1;	Length 1161;
Best Local Similarity	50.08;	Prod No 0 39;		

Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0

430 TGGTGTGGGCTTCGTCGAAGGCAAGGACGAGCAAGGCCGGCTGCTGCGGGCCGACGCTCA 489

Db 257 TGTGCGCCGCTCTGCTACTCCGAGGTCAGGAGGCGCGCGCTGCTGAGACCCCGCC 316
QY 490 TCCGCTAGCCCAACCTGGGCAACGCTCATCTCTCGGAGCGTCAGACC 539
Db 317 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACACCGCTCCATC 366

RESULT 5

US-08-475-742-3
Sequence 3, Application US/08475742
Patent No. 6121015

GENERAL INFORMATION:
APPLICANT: O'Malley, Karen L

APPLICANT: Todd, Richard D

TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor

FILE REFERENCE: WU 102 CON DIV

CURRENT APPLICATION NUMBER: US/08/475,742

CURRENT FILING DATE: 1995-06-07

EARLIER FILING DATE: 1994-06-16

EARLIER APPLICATION NUMBER: US 08/014,013

EARLIER FILING DATE: 1993-01-28

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 1367

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc.feature

LOCATION: (1)..(1367)

OTHER INFORMATION: D4 Dopamine Receptor cDNA

PUBLICATION INFORMATION:
AUTHORS: Van Tol, H. H.

AUTHORS: Bunzow, J. R.

TITLE: Cloning of the gene for a human dopamine D4 receptor

JOURNAL: Nature

VOLUME: 360

PAGES: 610-614

DATE: 1991

US-08-475-742-3

Query Match 1.8%; Score 41.2; DB 3; Length 1367;
Best Local Similarity 60.9%; Pred. No. 0.32;

Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTCGGGCTCTGCTGAGGCAAGGAGCGCAAGCGCGCTGCTGGGCGCAGCTCA 489
Db 363 TGTGTCGGGCTCTGCTGAGGCAAGGAGCGCGCTGCTGGGCGCAGCTCA 422

QY 490 TCCGCTAGCCCAACCTGGGCAACGCTCATCTCTCGGAGCGTCAGACC 539
Db 423 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACACCGCTCCATC 472

US-08-475-742-3

RESULT 6

US-08-056-051-1

Sequence 1, Application US/08056051

Patent No. 5516683

GENERAL INFORMATION:
APPLICANT: Grandy, David K

APPLICANT: Bunzow, James R

APPLICANT: Civelli, Olivier

APPLICANT: Van Tol, Hubert H.-M.

TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,051

FILING DATE: 19930429

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: No. 5516683man, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-C

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1370 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: 5'UTR

LOCATION: 1..103

FEATURE:
NAME/KEY: 3'UTR

LOCATION: 1268..1370

FEATURE:
NAME/KEY: CDS

LOCATION: 104..1267

US-08-056-051-1

Query Match 1.8%; Score 41.2; DB 1; Length 1370;
Best Local Similarity 60.9%; Pred. No. 0.32;

Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTCGGGCTCTGCTGAGGCAAGGAGCGCAAGCGCGCTGCTGGGCGCAGCTCA 489
Db 363 TGTGTCGGGCTCTGCTGAGGCAAGGAGCGCGCTGCTGGGCGCAGCTCA 422

QY 490 TCCGCTAGCCCAACCTGGGCAACGCTCATCTCTCGGAGCGTCAGACC 539
Db 423 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACACCGCTCCATC 472

US-07-928-611-17

Sequence 17, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civelli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611

ATTORNEY/AGENT INFORMATION:
NAME: NO. 5883226nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002

```

: INFORMATION FOR SEQ ID NO: 17
:
: SEQUENCE CHARACTERISTICS:
:
:   LENGTH: 1370 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA
:   FEATURE:

```


NAME: No. 588325nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1364..1466
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1363
US-08-487-811A-19

Query Match 1.8%; Score 41.2; DB 2; Length 1466;
Best Local Similarity 60.9%; Pred. No. 0.33;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTGGGCTTGTGCGAAGGAGGAGGAGGCGGCTGTGGGGCGGAGCTCA 489
DB 363 TGTGTGGGCTTGTGCTACTCTCCAGGTCAGGGTGGGCGGCTGTGAGCCCCCGCC 422

QY 490 TCCGCTACGCCAACCTGGGCAACGTCATCTCGGCAAGCGCTCAGCACC 539
DB 423 TGTGTGGGCGCTCTCATGTGCGCATGAGCTCATGCTGTGCAACCGCTTCATC 472

RESULT 15
US-09-060-694-19
Sequence 19, Application US/09060694
Patent No. 6203998
GENERAL INFORMATION:
APPLICANT: Clivell, Olivier
APPLICANT: Van Tol, Hubert H.M.
TITLE OF INVENTION: A No. 6203998e1 Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,694
FILING DATE: 15-APR-1998
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 6203998nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-MM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:

LENGTH: 1466 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1364..1466
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1363
US-09-060-694-19

Query Match 1.8%; Score 41.2; DB 4; Length 1466;
Best Local Similarity 60.9%; Pred. No. 0.33;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTGGGCTTGTGCGAAGGAGGAGGAGGCGGCTGTGGGGCGGAGCTCA 489
DB 363 TGTGTGGGCTTGTGCTACTCTCCAGGTCAGGGTGGGCGGCTGTGAGCCCCCGCC 422

QY 490 TCCGCTACGCCAACCTGGGCAACGTCATCTCGGCAAGCGCTCAGCACC 539
DB 423 TGTGTGGGCGCTCTCATGTGCGCATGAGCTCATGCTGTGCAACCGCTTCATC 472

Search completed: June 8, 2003, 18:16:29
Job time : 92.8557 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:51:03 ; Search time 2950 Seconds
(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229
Sequence: 1 cagggagctccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1793	80.4	2453	11	BC015220 Homo sapi
2	777.6	34.9	963	14	B0879880 AGENCOURT
3	754.2	33.8	1699	11	AK006549 Mus muscu
4	640.4	28.7	666	13	BI480798 H2RPE-043
5	582.4	26.1	592	13	BM663028 UT-E-C10-
6	577	25.9	773	10	BE410951 601303662

7	575	25.8	585	14	BM707948	BM707948	UT-E-C11-
8	565.2	25.4	593	10	BE385296	BE385296	60127572
9	528.8	23.7	735	13	BI756228	BI756228	603024265
10	525.8	23.6	537	14	BM691456	BM691456	UT-E-C11-
11	522	23.4	522	14	BM707649	BM707649	UT-E-C11-
12	508	22.8	525	14	BM685501	BM685501	UT-E-C10-
13	492.4	22.1	930	14	BQ436824	BQ436824	AGENCOURT
14	485.2	21.8	566	14	BM718338	BM718338	UT-E-B01-
15	475.4	21.3	503	14	BM685122	BM685122	UT-E-E01-
16	470.8	21.1	508	14	BM932117	BM932117	UT-E-E01-
17	470.8	20.9	474	9	AI190190	AI190190	qd37c03.x
18	466.8	20.9	469	9	AA307119	AA307119	EST178031
19	449.2	20.2	455	12	BG013943	BG013943	TI5-GN024
20	444.6	19.9	601	13	BG951790	BG951790	MR1-C7073
21	443	19.9	481	13	BM662514	BM662514	UT-E-C11-
22	437.4	19.6	531	14	BM685396	BM685396	UT-E-C10-
23	427.4	19.2	484	14	N31453	N31453	YX55408.r1
24	405.4	18.2	419	14	BM718146	BM718146	UT-E-E00-
25	401.4	18.0	508	10	BE236901	BE236901	145906 MA
26	391	17.5	766	10	BE275846	BE275846	601121656
27	379.8	17.0	1067	13	BM562042	BM562042	AGENCOURT
28	378.4	17.0	381	9	AI208843	AI208843	9938912.x
29	365.8	16.4	371	14	BM694530	BM694530	UT-E-C11-
30	360.6	16.2	603	9	AA205892	AA205892	zq47a06.r
31	354.8	15.9	793	13	BG961794	BG961794	602826439
32	347.2	15.6	444	14	N33227	N33227	Y707402.s1
33	347.2	15.3	514	13	BI343182	BI343182	371242 MA
34	307.4	13.8	599	13	BG951382	BG951382	MR1-C7073
35	300	13.5	525	14	BQ345562	BQ345562	PM3-NT031
36	299.4	13.4	317	14	N25339	N25339	YX55408.s1
37	293.4	13.2	305	9	AA318352	AA318352	EST20422
38	291.6	13.1	654	10	BE679896	BE679896	d58a11.y
39	283.4	12.7	732	13	BG965806	BG965806	602829577
40	282.8	12.7	297	9	AA991578	AA991578	OS60C10.s
41	271.6	12.2	394	13	BG950641	BG950641	MR1-C7073
42	268.8	12.1	566	13	BI443895	BI443895	da196b11.
43	246.2	11.0	271	9	AA317489	AA317489	EST19416
44	244.4	11.0	613	9	AI654606	AI654606	AL654606
45	243.4	10.9	422	10	AM594705	AM594705	h963d05.x

ALIGNMENTS

RESULT 1	BC015220	2453 bp	mRNA	linear	HTC 25-JUL-2002
LOCUS	BC015220				
DEFINITION	Homo sapiens, clone IMAGE:3877806, mRNA.				
ACCESSION	BC015220				
VERSION	BC015220.1	GI:21955361			
KEYWORDS	HTC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 2453)				
JOURNAL	Strausberg, R.				
	Direct Submission				
	Submitted (01-OCT-2001) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				

REMARK
COMMENT
Contact: MGC help desk
Email: cgaabbs-r@mail.nih.gov
Tissue Procurement: DCTP/DTP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
 Series: IRAC Plate: 14 Row: b Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

FEATURES
 source
 location/Qualifiers
 1..2453

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3877806"
 /tissue_type="lung, large cell carcinoma"
 /clone_id="N1H_MGC_68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

BASE COUNT 633 a 711 c 606 g 503 t
 ORIGIN

Query Match 80.4% Score 1793; DB 11; Length 2453;
 Best Local Similarity 95.6% Pred. No. 1.1e-289;
 Matches 1887; Conservative 0; Mismatches 5; Indels 81; Gaps 1;

257 GCTGCGCCCTCAGAGAAACAACACCTGATGTTGAGAAACTGACTGTGATGCGACAG 316
 557 GCTGCGCCCTCAGAGAAACAACACCTGATGTTGAGAAACTGACTGTGATGCGACAG 616
 317 CTACATCCAGCTCATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376
 617 CTACATCCAGCTCATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
 377 CCGCTGCTGAGAACAGTACAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
 677 CCGCTGCTGAGAACAGTACAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
 437 GGGCTTCTGAG 496
 737 GGGCTTCTGAG 796
 497 GGGCAACCTGGGCAACCTGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
 797 GGGCAACCTGGGCAACCTGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
 557 CCCGAGCGCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
 857 CCCGAGCGCCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
 617 GGGAAACTGAGAGCTTACACACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
 917 GGGAAACTGAGAGCTTACACACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
 677 GTCATGAAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
 977 GTCATGAAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
 737 GAGAGAGATGAAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
 1037 GAGAGAGATGAAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
 797 TATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
 1097 TATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
 857 TCTAGTGGGCGGAGCTTCTGAAACCCAGGCTTACCTGCGCATGAGCTGAGCT 916
 1157 TCTAGTGGGCGGAGCTTCTGAAACCCAGGCTTACCTGCGCATGAGCTGAGCT 1216
 917 CGTTGTCGCCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
 1217 CGTTGTCGCCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269

977 AGAGACATCATCAACCCCTTGGAGAGATGATGATGATTTTGAACCAACTGAGATTGT 1036
 1270 ----- 1269
 1037 CGACAGAAATTTGAGAGTGTCTCTGTTGGCTGTGTGATGATGATGATGATGATGATGAT 1096
 1270 -----AGGTGTCCTGTTGGCTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1315
 1097 GATGAGCCCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1156
 1316 GATGAGCCCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
 1157 CGCCAGTTCCTGAGAGCTTCTTATGAGGCTTCCACCTTCAACATGACCTGTAACAAAGA 1216
 1376 CGCCAGTTCCTGAGAGCTTCTTATGAGGCTTCCACCTTCAACATGACCTGTAACAAAGA 1435
 1217 GAGATGAGATTCAGACCCCAATCAGAGAGAGAGAGAGATGCTCAACGCTGGCATCATTTGG 1276
 1436 GAGATGAGATTCAGACCCCAATCAGAGAGAGAGAGAGATGCTCAACGCTGGCATCATTTGG 1495
 1277 CCGCTTCTGAGAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1336
 1496 CCGCTTCTGAGAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1555
 1337 ACTACTGTGGCCCAAGAGGATCCCTTCTCAGAGAGGCTTCCCAAAAACACAGAGC 1396
 1556 ACTACTGTGGCCCAAGAGGATCCCTTCTCAGAGAGGCTTCCCAAAAACACAGAGC 1615
 1397 AGCCAAACAGAACTTAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1456
 1616 AGCCAAACAGAACTTAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
 1457 GCGCTTCAAGTGTGGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1516
 1676 GCGCTTCAAGTGTGGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1735
 1517 CCTCAGCCCACTCCATGTTCTTCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1576
 1736 CCTCAGCCCACTCCATGTTCTTCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1795
 1577 TGTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1636
 1796 TGTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1855
 1637 AGTTTGAATGCTCTCAGAT 1696
 1856 AGTTTGAATGCTCTCAGAT 1915
 1697 AGTGAT 1756
 1916 AGTGAT 1975
 1757 CCTCAAGAACTTGTGAT 1816
 1976 CCTCAAGAACTTGTGAT 2035
 1817 GATTCCTTATGAGGCTTGTGAT 1876
 2036 GATTCCTTATGAGGCTTGTGAT 2095
 1877 GGGATGCTTCCAGAT 1936
 2096 GGGATGCTTCCAGAT 2155
 1937 ACAGCATACAGTGTCCAGAT 1996
 2156 ACAGCATACAGTGTCCAGAT 2215
 1997 AGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
 2216 AGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2275
 2057 TGTGAAGAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2116

DB 2276 TGTGAAGCTAGACGACCATTTGAAACATTAACTGACTGTGATTCAGAGTCCGG 2335
OY 2117 AACCTTAGTCTATCTGAAATCCAGACGACCACTAGTATATCTGCCAACTAAG 2176
DB 2336 AACCTTAGTCTATCTGAAATCCAGACGACCACTAGTATATCTGCCAACTAAG 2395
OY 2177 AGTTAATATATACAAATACCTGTTAAAAA 2229
DB 2396 AGTTAATATATACAAATACCTGTTAAAAA 2448
RESULT 2
BQ879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT 8241531 lupski_dorsal.root.ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6180559 5', mRNA sequence.
ACCESSION BQ879880
VERSION BQ879880.1 GI:22271888
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13563 row: 1 column: 08
High quality sequence stop: 623.
FEATURES
Source location/Qualifiers
1..963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORE6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 205 a 289 c 265 g 202 t 2 others
ORIGIN
Query Match 34.9%; Score 777.6; DB 14; Length 963;
Best Local Similarity 96.2%; Pred. No. 3.2e-120;
Matches 795; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 445 TCGAAGGACGAGGACGAGGAGCGGCTGCTGCGGCGACGCTATCCGCTAGCCCAAC 504
DB 26 TCGAAGGACGAGGACGAGGAGCGGCTGCTGCGGCGACGCTATCCGCTAGCCCAAC 85
OY 505 TGGGCAACGTCGTCATCTCGGCGAGCGTCAAGCAGCGAGTCTACAGCGCTTCCCGACG 564
DB 86 TGGGCAACGTCGTCATCTCGGCGAGCGTCAAGCAGCGAGTCTACAGCGCTTCCCGACG 145

OY 565 CCCAGCAGCTGTGCAAGAGCGCTTATGACTCCGCGACAAACAAAGCACTTGGAGAAC 624
DB 146 CCCAGCAGCTGTGCAAGAGCGCTTATGACTCCGCGACAAACAAAGCACTTGGAGAAC 205
OY 625 TGAGCTACCAACAAATGTTCTGGGTGCGCTGGGTGGTGGTGGTGGTGGTGGTGGT 684
DB 206 TGAGCTACCAACAAATGTTCTGGGTGCGCTGGGTGGTGGTGGTGGTGGTGGTGGT 265
OY 685 AGCGGTGCTGGAGGTGCAATCCGGAGCCATCTGCTCCAGAGCCTGCTGAAGAGA 744
DB 266 AGCGGTGCTGGAGGTGCAATCCGGAGCCATCTGCTCCAGAGCCTGCTGAAGAGA 325
OY 745 TGAACACCTTGGCTACTGAGTGTGACACCTGTATCCCTACAGCTGGATATATCCAC 804
DB 326 TGAACACCTTGGCTACTGAGTGTGACACCTGTATCCCTACAGCTGGATATATCCAC 385
OY 805 TGGTGTATACAGAGTGTGACCTGGCGGTGTACAGCTTCTTCTGACTGTCTAGTTG 864
DB 386 TGGTGTATACAGAGTGTGACCTGGCGGTGTACAGCTTCTTCTGACTGTCTAGTTG 445
OY 865 GGGGCGAGTTTGAACCCAGCCAGCCCTACCTGAGCTGAGCTGCTGCTGCTGCTGCT 924
DB 446 GGGGCGAGTTTGAACCCAGCCAGCCCTACCTGAGCTGAGCTGCTGCTGCTGCTGCT 505
OY 925 CGCTCTTACAGTCTGCAAGTCTTCTTATGTTGGCTGCTGAAGTGGACAGCAGC 984
DB 506 CGCTCTTACAGTCTGCAAGTCTTCTTATGTTGGCTGCTGAAGTGGACAGCAGC 565
OY 985 TCATCAACCCCTTTGAGAGAGATGATGATTTTGAACCACTGGATGTGACAGCA 1044
DB 566 TCATCAACCCCTTTGAGAGAGATGATGATTTTGAACCACTGGATGTGACAGCA 625
OY 1045 ATTGAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
DB 626 ATTGAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 685
OY 1105 CGGACATGTAAGTAATTAAGCCGAGACAGCCGCTTACAGAGCTGCTTCCGCCAGT 1164
DB 686 CGGACATGTAAGTAATTAAGCCGAGACAGCCGCTTACAGAGCTGCTTCCGCCAGT 745
OY 1165 TCCGTGAGGCTCTCTTATGAGGCTCCACCTTCAACATCAGCTGAAAGAGAGATGG 1224
DB 746 TCCGTGAGGCTCTCTTATGAGGCTCCACCTTCAACATCAGCTGAAAGAGAGATGG 805
OY 1225 AGTTCAGCCCAATACGAGAGGAGGAGGATGCTCAGCTGCGCAT 1270
DB 806 AATTCAGCCCAATACGAGAGGAGGAGGATGCTCAGCTGCGCAT 851
RESULT 3
AK006549 1699 bp mRNA linear HTC 19-JAN-2002
LOCUS AK006549
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700030H21; vitellogenin macular dystrophy 2 homolog
(human), full insert sequence.
ACCESSION AK006549
VERSION AK006549.1 GI:12839710
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male testis cDNA to mRNA,
clone:1700030H21.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

QY 873 TTTTGAACCCAGCCAGCCCTACCTGGCCTGAGCTGACCTCTGTGGCCCTTC 932
 DB 600 TTTTGAACCCAGCCAGCCCTACCTGGCCTGAGCTGACCTCTGTGGCCCTTC 659
 QY 933 AGCTCTGCAAGTCTTCTATGATGGCTGGCTGAGAGTGGCAGAGAGCTCACTAC 992
 DB 660 ACAATCTGCAATTTTATTTCAATGGCTGGCTGGAGTGGCAGAGAGCTCACTAC 719
 QY 993 CCCTTGGAGAGATGATGATGATTTTGGAGCACTGATTTGTCAGCAATTTTGG 1052
 DB 720 CCCTTGGAGAGAGATGATGATTTTGGAGCACTGATTTGTCAGCAATTTTGG 779
 QY 1053 GTGTCCCTGTTGGCTGTGGAGTATGCAACAGAGCTGCTGGATGGAGCCGAGC 1112
 DB 780 GTGTCCCTGTTGGCTGTGGAGTATGCAACAGAGCTGCTGGATGGAGCCGAGC 839
 QY 1113 TACTGATTAAGCCGAGCAGCCCTACAGCCCTGCTGCTGCTGCTGCTGCTGCA 1172
 DB 840 TACTGATTAAGCCGAGCAGCCCTACAGCCCTGCTGCTGCTGCTGCTGCTGCA 899
 QY 1173 GCTCTTTTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1232
 DB 900 CATCTTTCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
 QY 1233 CCAATTCAGAGG-----ACGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1280
 DB 960 TCAAAAGAGAGGCTGACACGGATTAAGAGAGTGGCTATAGCAGCAGATGCTGCT 1019
 QY 1281 TTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340
 DB 1020 TTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
 QY 1341 CTGTGGCCAGAGAGGAAATCCCTTCTGCAAGAGGCTGCTGCTGCTGCTGCTGCT 1400
 DB 1080 TTGTGTCTTAAGA-----ACCCCTCTGCAAGGCTGCTGCTGCTGCTGCTGCTG 1133
 QY 1401 AAACAGAACTTATGAGGCTGCAAGAGCAACAGGCTGCTGCTGCTGCTGCTGCTG 1460
 DB 1134 AAACAGAACTTATGAGGCTGCAAGAGCAACAGGCTGCTGCTGCTGCTGCTGCTG 1172
 QY 1461 TTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1520
 DB 1173 TTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1232
 QY 1521 AGCCCACTCC 1531
 DB 1233 AGCCCACTCC 1243

RESULT 4
 LOCUS B1480798 666 bp mRNA linear EST 28-FEB-2002
 DEFINITION H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
 5, similar to vitelliform macular dystrophy (Best disease, mRNA
 sequence.
 ACCESSION B1480798
 VERSION B1480798.1 GI:18998607
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 666)
 AUTHORS Buraczynska, M., Mears, A. J., Zarepari, S., Farjo, R., Filipova, E., Yuan, Y., Macnee, S. P., Hughes, B., and Swarcop, A.
 TITLE Towards an expression profile of native human retinal pigment epithelium: Identification of a non-redundant set of more than 1100 genes
 JOURNAL Unpublished (2001)
 COMMENT Contact: Swarcop, A.
 Department of Ophthalmology and Visual Sciences
 Kellogg Eye Center, University of Michigan

540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
 Tel: 734 615 2246
 Fax: 734 647 0228
 Email: swarcop@umich.edu
 PCR Primers
 FORWARD: M13/PUC-Reverse - ccagatcagagctgttaaacg
 BACKWARD: M13/PUC-Forward - agcgataacattcacacag
 Seq primer: M13/PUC-Reverse.
 Location/Qualifiers
 source
 1. 666
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Human Retinal Pigment Epithelium (2)"
 /issue_type="Native Retinal Pigment Epithelium sheets"
 /dev_stage="juvenile"
 /note="Organ: Retina; Vector: pSPori1"

BASE COUNT 133 a 224 c 171 g 136 t 2 others
 ORIGIN

Query Match 28.7%; Score 640.4; DB 13; Length 666;
 Best Local Similarity 99.4%; Pred. No. 2.7e-97;
 Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1.

QY 10 CCACGAGCTAGTCCGACAGCTTCTGTGGATCATGGAACCACTGGAAACCCACTG 69
 DB 12 CCACGAGCTAGTCCGACAGCTTCTGTGGATCATGGAACCACTGGAAACCCACTG 71
 QY 70 ACCAAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
 DB 72 ACCAAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
 QY 130 TGGCTAATGCGCGCTTATAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
 DB 132 TGGCTAATGCGCGCTTATAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
 QY 190 ACAAGCTGATATATGCGAGTCTTATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 249
 DB 192 ACAAGCTGATATATGCGAGTCTTATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 251
 QY 250 TTTATAGGCTGCGGCTCAGAGGAAGCAACAGCTGATGTTGGAACCTGCTGATTT 309
 DB 252 TTTATAGGCTGCGGCTCAGAGGAAGCAACAGCTGATGTTGGAACCTGCTGATTT 311
 QY 310 GCGACAGTACATCAGCTCATCCCATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 369
 DB 312 GCGACAGTACATCAGCTCATCCCATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 371
 QY 370 TCGTACCCGCTGTGTGAACCACTACGAGAACTGCCGTGGCCGACCGCTCATGAGCC 429
 DB 372 TCGTACCCGCTGTGTGAACCACTACGAGAACTGCCGTGGCCGACCGCTCATGAGCC 431
 QY 430 TGGTGTGCGGCTGCTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 489
 DB 432 TGGTGTGCGGCTGCTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 491
 QY 490 TCGCTACGCAACTGTGGGCAAGCTGCTATCTTCTGCGCAGCTGACGACGCACTTACA 549
 DB 492 TCGCTACGCAACTGTGGGCAAGCTGCTATCTTCTGCGCAGCTGACGACGCACTTACA 551
 QY 550 AGCGTTCGCCAGCGCCGACACCTGTGTGCAAGAGCTTATGATCTCCGAGACACACA 609
 DB 552 AGCGTTCGCCAGCGCCGACACCTGTGTGCAAGAGCTTATGATCTCCGAGACACACA 610
 QY 610 AGAGTGTGGAACACGAGCTTACACACAACTGTTCTGGGTGCGCTGGGTGG 665
 DB 611 AGAGTGTGGAACACGAGCTTACACACAACTGTTCTGGGTGCGCTGGGTGG 666

RESULT 5
 LOCUS BM663028/c 592 bp mRNA linear EST 27-FEB-2002
 DEFINITION UI-E-C10-aad-h-10-0-UI.s1 UI-E-C10 Homo sapiens cDNA clone
 UI-E-C10-aad-h-10-0-UI 3', mRNA sequence.

ACCESSION BM663028
 VERSION BM663028.1 GI:18968017
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 592)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLYA-Yes.

FEATURES
 source Location/Qualifiers
 1..592
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UT-E-C10-aad-h-10-0-UT"
 /clone_lib="UT-E-C10"
 /clone_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UT-E-C10 is a CDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand CDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded CDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand CDNA contains a library tag sequence that is located between the Not I site and the (GT)₁₈ tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_LIB-UT-E-C10
 TAG_TISSUE-RPE and Choroid
 TAG_SEQ-ACCTA"

BASE COUNT 133 a 111 c 142 g 206 t
 ORIGIN
 Query Match 26.1%; Score 582.4; DB 13; Length 592;
 Best Local Similarity 99.0%; Pred. No. 1.3e-87;
 Matches 586; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1627 GGGCCAGAAAAGTTTGAATTCCTCTCGAGAGCGAGGCGCTTGATGAGACCCAG 1686
 |||
 DB 592 GGGCCAGAAAAGTTTGAATTCCTCTCGAGAGCGAGTGGGCGCTTGATGAGACCCAG 533
 |||
 QY 1687 AAGATCTCAAGTGAAGGAGGAGAAAGTGTGAAGTTAACTGACGAGATGCCAGATCC 1746
 |||
 DB 532 AAGATCTCAAGTGAAGGAGGAGAAAGTGTGAAGTTAACTGACGAGATGCCAGATCC 473
 |||
 QY 1747 CCGAAATACCTCAAGAACCTTTGAGAACATACCAACCAACATACACATCACTCA 1806
 |||
 DB 472 CCGAAATACCTCAAGAACCTTTGAGAACATACCAACCAACATACACATCACTCA 413

QY 1807 AAGATCAATGATCTTATTTGGCCTTGAGAAAACAGGAGTGAACATTTCTTAACCTG 1866
 |||
 DB 412 AAGATCAATGATCTTATTTGGCCTTGAGAAAACAGGAGTGAACATTTCTTAACCTG 353
 |||
 QY 1867 CTTCCTTAAGGGAGATGCTTCGCGAGCAGGCTCCTCAGCTGTGTATACACAGAGACAC 1926
 |||
 DB 352 CTTCCTTAAGGGAGATGCTTCGCGAGCAGGCTCCTCAGCTGTGTATACACAGAGACAC 293
 |||
 QY 1927 TGATCCAGTACAGCCATACAGCTGTCCACACTGAAAGAACGTCTCTAACAAGCTGTA 1986
 |||
 DB 292 TGATCCAGTACAGCCATACAGCTGTCCACACTGAAAGAACGTCTCTAACAAGCTGTA 233
 |||
 QY 1987 TCAATGCTTACCTTAT 2046
 |||
 DB 232 TCAATGCTTACCTTAT 173
 |||
 QY 2047 TCATTAATAAGTGTGAAGCTAGACTGAACCATTTGAAACATTTAACTCAGACTGTGAT 2106
 |||
 DB 172 TCATTAATAAGTGTGAAGCTAGACTGAACCATTTGAAACATTTAACTCAGACTGTGAT 113
 |||
 QY 2107 CAGAGTGGGAAACCTTATGTTCTATCTGATTCAGATCCAGACCCACACTTATATCTGCC 2166
 |||
 DB 112 CAGAGTGGGAAACCTTATGTTCTATCTGATTCAGATCCAGACCCACACTTATATCTGCC 53
 |||
 QY 2167 CAACCTATGAGTTTAAATATATACAAATCTCTTAAATATATATATATATATATATAT 2218
 |||
 DB 52 CAACCTATGAGTTTAAATATATACAAATCTCTTAAATATATATATATATATATATAT 1

RESULT 6
 BE410951 773 bp mRNA linear EST 21-JUL-2000
 LOCUS 601303662p1 NIH_MGC_21 Homo sapiens CDNA clone IMAGE:3638175 5',
 DEFINITION mRNA sequence.
 BE410951
 BE410951.1 GI:9347401
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 773)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabps-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LHCW338 row: 1 column: 16
 High quality sequence stop: 662.

FEATURES
 source Location/Qualifiers
 1..773
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3638175"
 /clone_lib="NIH_MGC_21"
 /tissue_type="Choriocarcinoma"
 /note="Organ: placenta; Vector: pOT87; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 204 a 239 c 176 g 154 t

ORIGIN

Query Match 25.98; Score 577; DB 10; Length 773;

Best Local Similarity 99.08; Pred. No. 9.6e-87; Matches 591; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

1253 GGATGCTACGGCTGATCATGTTGGCGCTTCTTGGCGCTGAGCCATGATGATCAGCAGTCC 1312
 1 GGATGCTACGGCTGATCATGTTGGCGCTTCTTGGCGCTGAGCCATGATGATCAGCAGTCC 60

1313 TCCGAGGCAAACTCAAGAGCAAACTATGTTGGCCCAAGAGGGAATCCCTTTCACAGA 1372
 61 TCCGAGGCAAACTCAAGAGCAAACTATGTTGGCCCAAGAGGGAATCCCTTTCACAGA 120

1373 GGGCTTCCC-AAAACCAAGAGCAGCAGCAAAAGAGCTTGGGCGCAGAGAACACACA 1431
 121 GGGCTTCCC-AAAACCAAGAGCAGCAGCAAAAGAGCTTGGGCGCAGAGAACACACA 180

1432 AGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCCACTGATCAGAGCCAG 1491
 181 AGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCCACTGATCAGAGCCAG 240

1492 GCTACTACAGTGCACACAGAGCGCCCTCAGCGCCCACTCCATGTTCTTCCCTAGAAC 1551
 241 GCTACTACAGTGCACACAGAGCGCCCTCAGCGCCCACTCCATGTTCTTCCCTAGAAC 300

1552 CATCAGGCGGCTCAAAAGCTTACAGTGTGACAGAGCTGACAGCAAAAGAGCTTAA 1611
 301 CATCAGGCGGCTCAAAAGCTTACAGTGTGACAGAGCTGACAGCAAAAGAGCTTAA 360

1612 AGACTGTAGTGTGAGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCGCT 1671
 361 AGACTGTAGTGTGAGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCGCT 420

1672 TGATGAGACACCCAGAGATATCTCAGTGTGAGAGAGAGAAAGTGTGAGTTTAACTGAGCG 1731
 421 TGATGAGACACCCAGAGATATCTCAGTGTGAGAGAGAGAAAGTGTGAGTTTAACTGAGCG 480

1732 ATATGCGCAGATGATGCGCAAAATGCTCAAGCAAAAGCTTGAACATCACCACCAACA 1791
 481 ATATGCGCAGATGATGCGCAAAATGCTCAAGCAAAAGCTTGAACATCACCACCAACA 540

1792 TACACACTACACTCAAAAGATCATGATGATCTTATTTGGGCGCTTGAAGAAAGAGATG 1848
 541 TACACACTACACTCAAAAGATCATGATGATCTTATTTGGGCGCTTGAAGAAAGAGATG 597

RESULT 7 585 bp mRNA linear EST 28-FEB-2002
 BM707948
 LOCUS UI-E-C11-aft-g-11-0-UI.1 UI-E-C11 Homo sapiens cDNA clone
 DEFINITION UI-E-C11-aft-g-11-0-UI.1 mRNA sequence.
 BM707948
 ACCESSION BM707948.1 GI:19021206
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 585)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@iuii.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
 1..585
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-aft-g-11-0-UI"
 /clone_1b="UI-E-C11"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pRT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (d)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

BASE COUNT 187 a 145 c 121 g 130 t 2 others

ORIGIN

Query Match 25.88; Score 575; DB 14; Length 585;
 Best Local Similarity 98.88; Pred. No. 2.3e-86;
 Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1556 AGCGCGCTCAAAAGCTTCACTGTGACAGGAGTATGACACCAAGAGCAAAAGCTTAAAGC 1615
 1 AGCGCGCTCAAAAGCTTCACTGTGACAGGAGTATGACACCAAGAGCAAAAGCTTAAAGC 60

1616 TGTGAGTCTGTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTGAT 1675
 61 TGTGAGTCTGTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTAT 120

1676 GGAGCACCAGAGATATCTCAAGTGTGAGAGAGAAAGCTGTGAGTTAACTGACGATAT 1735
 121 GGAGCACCAGAGATATCTCAAGTGTGAGAGAGAAAGCTGTGAGTTAACTGACGATAT 180

1736 GCCAGAGATCCCGGAAATACCTTCAAAAGAACTTTGGAAACATGACCAACCAACTATCA 1795
 181 GCCAGAGATCCCGGAAATACCTTCAAAAGAACTTTGGAAACATGACCAACCAACTATCA 240

1796 CACTACACTCAAAAGATCATGAGATGCTTATTTGGGCGCTTGAAGAAAGAGATGAAGACA 1855
 CACTACACTCAAAAGATCATGAGATGCTTATTTGGGCGCTTGAAGAAAGAGATGAAGACA 1855

241 CACTACACTCAAAAGATCATGAGATGCTTATTTGGGCGCTTGAAGAAAGAGATGAAGACA 300
 1856 TTCTTAACCTGCTTCTTAATGAGGAGTGTGCGCCAGCCAGAGTCTCACTGTGTACAC 1915

301 TTCTTAACCTGCTTCTTAATGAGGAGTGTGCGCCAGCCAGAGTCTCACTGTGTACAC 360

1916 CAGCAGACACTGATCAGTCAAGCAGCCATACAGCTGTGCACACTAAGAGCTGTCTAC 1975
 CAGCAGACACTGATCAGTCAAGCAGCCATACAGCTGTGTGCACACTAAGAGCTGTCTAC 1975

361 CAGCAGACACTGATCAGTCAAGCAGCCATACAGCTGTGTGCACACTAAGAGCTGTCTAC 420

1976 AACAGCCGGAATCAATGTTAGCTTAATGATTAATAATCCAGACTACTTACGCTTAA 2035
 AACAGCCGGAATCAATGTTAGCTTAATGATTAATAATCCAGACTACTTACGCTTAA 2035

421 AACAGCCGGAATCAATGTTAGCTTAATGATTAATAATCCAGACTACTTACGCTTAA 480

2036 ATGCTTTTATTCATTAATAAGTGTGAAGCTGAGTGAACATTTGAAACATTAATCA 2095
 ATGCTTTTATTCATTAATAAGTGTGAAGCTGAGTGAACATTTGAAACATTAATCA 2095

481 ATGCTTTTATTCATTAATAAGTGTGAAGCTGAGTGAACATTTGAAACATTAATCA 540

OY 2096 GACTGTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGATCA 2140
 DB 541 GACTGTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGATCA 585
 RESULT 8
 BE385296 593 bp mRNA linear EST 21-JUL-2000
 LOCUS 601277572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618665 5',
 DEFINITION mRNA sequence.
 ACCESSION BE385296
 VERSION BE385296.1 GI:9330661
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 593)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 cDNA Library Preparation: Ling Hong/Rubln Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov
 Plate: L1CM287 row: 0 column: 18
 High quality sequence stop: 593.
 Location/Qualifiers
 FEATURES
 source 1..593
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3618665"
 /clone_11b="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACAG(G). Size selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using Zap-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 176 a 177 c 132 g 108 t
 ORIGIN
 Query Match 25.4%; Score 565.2; DB 10; Length 593;
 Best Local Similarity 99.3%; Pred. No. 9.7e-85;
 Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 OY 1268 CATCATTTGGCGGCTTCTAGGCTGAGTCCCATGATACCATCTCCAGGCAATCTC 1327
 DB 1 CATCATTTGGCGGCTTCTAGGCTGAGTCCCATGATACCATCTCCAGGCAATCTC 60
 OY 1328 AAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTTCACAGGGGCTGCC-AAAA 1386
 DB 61 AAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTTCACAGGGGCTGCC-AAAA 120
 OY 1387 ACCACAAAGCAGCAACAAACAGAGTTAGGGGCCAGAGAGCAACAAAGGCTGAGAGTTA 1446
 DB 121 ACCACAAAGCAGCAACAAACAGAGTTAGGGGCCAGAGAGCAACAAAGGCTGAGAGTTA 180
 OY 1447 AGGCTGTGAGCGCTTCAAGTCTGCGCCACTGTATCAGAGCCAGGCTACTAGTCC 1506
 DB 181 AGGCTGTGAGCGCTTCAAGTCTGCGCCACTGTATCAGAGCCAGGCTACTAGTCC 240
 OY 1507 CACAGAGCGCCCTCAGGCCCACTCCATGTTCTTCCCTTGAACATCAGCGCCCTCAA 1566

DB 241 CACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACATCAGCGCCCTCAA 300
 OY 1567 AGCTTCACAGTGTACAGGCAATGACACCAAGCAAAAGCTTAAAGACTGTGATCTG 1626
 DB 301 AGCTTCACAGTGTACAGGCAATGACACCAAGCAAAAGCTTAAAGACTGTGATCTG 360
 OY 1627 GGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGTATGAGCAACCCAG 1686
 DB 361 GGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGTATGAGCAACCCAG 420
 OY 1687 AAGTATCTCAGTGTAGAGGAGAAAGTGTGGAGTTTAACTGACGATATGACAGATTC 1746
 DB 421 AAGTATCTCAGTGTAGAGGAGAAAGTGTGGAGTTTAACTGACGATATGACAGATTC 480
 OY 1747 CCGAAATACCTCAAGAACTTTGGAACAAATCACAACCAACATACACTACACTCA 1806
 DB 481 CCGAAATACCTCAAGAACTTTGGAACAAATCACAACCAACATACACTACACTCA 540
 OY 1807 AAGATCATGTGATCTTATTTGGGCTTGGAAAACAGGATG 1848
 DB 541 AAGATCATGTGATCTTATTTGGGCTTGGAAAACAGGATG 582

RESULT 9
 BI756228 735 bp mRNA linear EST 25-SEP-2001
 LOCUS 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5',
 DEFINITION mRNA sequence.
 ACCESSION BI756228
 VERSION BI756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://Image.lnl.gov
 Plate: L1AM11487 row: a column: 18
 High quality sequence stop: 577.
 Location/Qualifiers
 FEATURES
 source 1..735
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194649"
 /clone_11b="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
 Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH-MGC Library."
 BASE COUNT 135 a 229 c 214 g 157 t
 ORIGIN
 Query Match 23.7%; Score 528.8; DB 13; Length 735;
 Best Local Similarity 92.6%; Pred. No. 1.1e-78;
 Matches 613; Conservative 0; Mismatches 37; Indels 12; Gaps 5;

```

OY 257 GCTGGCCCTCAGGAGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 316
DB 59 GCTGGCCCTCAGGAGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 118
OY 317 CTATATCCAGCTCATCCCATTTCTTCCTGCTGCTGGGCTTCTACGTGACGTGCTGAC 376
DB 119 CTATATCCAGCTCATCCCATTTCTTCCTGCTGCTGGGCTTCTACGTGACGTGCTGAC 178
OY 377 CCGCTGTGTGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 436
DB 179 CCGCTGTGTGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 238
OY 437 GGGCTGTGTGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 496
DB 239 GGGCTGTGTGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 298
OY 497 CCGCAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 556
DB 299 CCGCAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 358
OY 557 CCGCAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 616
DB 359 CCGCAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 418
OY 617 GAGAGAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 676
DB 419 GAGAGAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 478
OY 677 GTCATATGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 736
DB 479 GTCATATGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 538
OY 737 GAGAGAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 794
DB 539 GAGAGAACTGAGGAGAACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 598
OY 795 AGTATCCCACTGATGTTGAGAACTGACTCTGTATGCGACAG 847
DB 599 AGTATCCCACTGATGTTGAGAACTGACTCTGTATGCGACAG 658
OY 848 CCGGAACTGATGTTGAGAACTGACTCTGTATGCGACAG 904
DB 659 CCGGAACTGATGTTGAGAACTGACTCTGTATGCGACAG 718
OY 905 TG 906
DB 719 TG 720

```

RESULT 10
LOCUS BM691456 537 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CII-abe-f-10-0-UI.1 UI-E-CII Homo sapiens cDNA clone
ACCESSION BM691456
VERSION BM691456.1 GI:19004714
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 537)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: mssoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..537
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-CII-abe-f-10-0-UI"
 /clone_1bp="UI-E-CII"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-CII is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pRT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCGA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

BASE COUNT 133 a 172 c 139 g 93 t
ORIGIN

Query Match 23.6%; Score 525.8; DB 14; Length 537;
 Best Local Similarity 99.6%; Pred. No. 3,7e-78;
 Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1020 GAGACCACTGGATGTCGACAGAAATTTGACAGTGTCCCTGTTGGCTGTGATGATG 1079
DB 8 GAGACCACTGGATGTCGACAGAAATTTGACAGTGTCCCTGTTGGCTGTGATGATG 67
OY 1080 CACCAAGACCTGCTCGATGAGAGCCGAGCAATGTAATTAACCCGAGACACAGCC 1139
DB 68 CACCAAGACCTGCTCGATGAGAGCCGAGCAATGTAATTAACCCGAGACACAGCC 127
OY 1140 CCTACACAGCTGCTTCCGCGCAGTTCGTCGAGCCCTCTTATGAGGCTCCACTTCAAC 1199
DB 128 CCTACACAGCTGCTTCCGCGCAGTTCGTCGAGCCCTCTTATGAGGCTCCACTTCAAC 187
OY 1200 ATCAGCTGTAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGATGCT 1259
DB 188 ATCAGCTGTAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGATGCT 247
OY 1260 CAGCTGATATGATGGCGGCTTCTTAAGGCTGTCAGTCCATGATCACCATCTCCAGG 1319
DB 248 CAGCTGATATGATGGCGGCTTCTTAAGGCTGTCAGTCCATGATCACCATCTCCAGG 307
OY 1320 GCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCAGAGGCTG 1379
DB 308 GCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCAGAGGCTG 367
OY 1380 CCAAAACACAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCAGAGGCTG 1439
DB 368 CCAAAACACAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCAGAGGCTG 427
OY 1440 AAGCTTAAGCTGTGAGGCTTCAAGTGTGAGGCTTCAAGTGTGAGGCTTCAAGTGTGAG 1499
DB 428 AAGCTTAAGCTGTGAGGCTTCAAGTGTGAGGCTTCAAGTGTGAGGCTTCAAGTGTGAG 487
OY 1500 AGTGCCCAAGAGAGCCCTCAGGCCCACTCCCACTGTTCTTCCCTAG 1548

```

Db 488 AGTGCCACAGAGCCCTCAGCCCTCAGCTGATGTTCTCCCGCTAG 536

|||||

RESULT 11
BM707649 522 bp mRNA linear EST 28-FEB-2002

LOCUS
DEFINITION
UI-E-CII-afs-o-15-0-UI.r1 UI-E-CII Homo sapiens cDNA clone
UI-E-CII-afs-o-15-0-UI 5', mRNA sequence.

ACCESSION
BM707649

VERSION
BM707649.1 GI:19020907

KEYWORDS
EST.

SOURCE
human.

ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
Normalisation and subtraction: two approaches to facilitate gene discovery

JOURNAL
MEDLINE
COMMENT
Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

FEATURES
source
Location/Qualifiers
1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-CII-afs-o-15-0-UI"
/clone_1lb="UI-E-CII"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site.1: EcoR I; Site.2: Not I; UI-E-CII is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (dT)18 tail. The sequence tag for this library is ACCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT
ORIGIN
100 a 182 c 130 g 110 t

Query Match. 23.4%; Score 522; DB 14; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.6e-77;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

46 CGGACCCACCTGAGACCCACCTGACCCACCTGTCGACGCCACCTGGCGCA 105
|||||

Db 1 CGGACCCACCTGAGACCCACCTGACCCACCTGTCGACGCCACCTGGCGCA 60

QY 106 TGACATCACTTACACAGCAAGTGGCTATGCCGCTTGAGCTCTTCCGCCCTGC 165

Db 61 TGACATCACTTACACAGCAAGTGGCTATGCCGCTTGAGCTCTTCCGCCCTGC 120

|||||

QY 166 TCGTGTGCTGGCGGCGAGCATCTACAGCTGCTATATGGCGATTCTTATCTTCG 225

Db 121 TCGTGTGCTGGCGGCGAGCATCTACAGCTGCTATATGGCGATTCTTATCTTCG 180

QY 226 TCGTGTGCTGTACATCATCCGCTTATTTATATAGCTGGCGCTCAGGAGAAACACAGCTGA 285

Db 181 TCGTGTGCTGTACATCATCCGCTTATTTATATAGCTGGCGCTCAGGAGAAACACAGCTGA 240

QY 286 TGTGTAGAAACGACTGTATTTGGAGAGCTATACAGCTCATCCCATTTCCCTGC 345

Db 241 TGTGTAGAAACGACTGTATTTGGAGAGCTATACAGCTCATCCCATTTCCCTGC 300

QY 346 TCGTGTGCTGTACATCATCCGCTTATTTATATAGCTGGCGCTCAGGAGAAACACAGCTGC 405

Db 301 TCGTGTGCTGTACATCATCCGCTTATTTATATAGCTGGCGCTCAGGAGAAACACAGCTGC 360

QY 406 CGTGGCGGACCGGCTCATGAGCTGTGCTGGGCTTCTCGAAGGCAAGAGACGACCAAG 465

Db 361 CGTGGCGGACCGGCTCATGAGCTGTGCTGGGCTTCTCGAAGGCAAGAGACGACCAAG 420

QY 466 GCCGCTGTGGGCGGCGAGCTCATCCGCTACAGCCAACTGGGCAAGGCTCATCTGC 525

Db 421 GCCGCTGTGGGCGGCGAGCTCATCCGCTACAGCCAACTGGGCAAGGCTCATCTGC 480

QY 526 GCAGGCTGACACACCGAGTCTACAAAGCGCTTCCCGAGCGCC 567

Db 481 GCAGGCTGACACACCGAGTCTACAAAGCGCTTCCCGAGCGCC 522

|||||

RESULT 12
BM685501 526 bp mRNA linear EST 27-FEB-2002

LOCUS
DEFINITION
UI-E-CIO-sad-h-10-0-UI.r1 UI-E-CIO Homo sapiens cDNA clone
UI-E-CIO-sad-h-10-0-UI 5', mRNA sequence.

ACCESSION
BM685501

VERSION
BM685501.1 GI:18995397

KEYWORDS
EST.

SOURCE
human.

ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
Normalisation and subtraction: two approaches to facilitate gene discovery

JOURNAL
MEDLINE
COMMENT
Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

FEATURES
source
Location/Qualifiers
1..526
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-CIO-sad-h-10-0-UI"
/clone_1lb="UI-E-CIO"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

RESULT 14
LOCUS BM718338
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

JOURNAL MEDLINE COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL MEDLINE COMMENT

BM718338 566 bp mRNA linear EST 01-MAR-2002
UT-E-BOL-a)a-j-23-0-UI.r1 UT-E-BOL Homo sapiens cDNA clone.
BM718338
BM718338 GI:19036537
EST.
human.
Homo sapiens
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
l (bases 1 to 566)
Bonaldo,A.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9365
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: MJ3 Reverse.

FEATURES	Location/Qualifiers
Source	1. .566

```

BASE COUNT      107 a      182 c      145 g      132 t
ORIGIN
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-E01-ajs-j-23-0-01"
/clone_11b="UI-E-E01"
/issue_type="fetal eye"
/issue_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-E01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGCATATACC. This library was created for the program, GeneDiscover in the Visual System, supported by National Eye Institute (NEI)."
```

Query Match	21.8%	Score 485.2	DB 14	Length 566
Best Local Similarity	99.4%	Pred. No. 2.2e-71		
Matches 487	Conservative 0	Mismatches 3	Indels 0	Gaps 0

Qy	18	TTTACTGCGAGAACTTTCGTGTGGATTCATCGAGACCACTGGAAACCCACCTGACCCAAAG	77
			1
Db	1	CTATCTGCGAGAACTTTCGTGTGGATTCATCGAGACCACTGGAAACCCACCTGACCCAAAG	60
Qy	78	CCAACTGCTGAGAGCCACTGGCTGGCATGACATCACTTACACAAGCAAGTGGCTTAAT	137
			1
Db	61	CCAACTGCTGAGAGCCACTGGCTGGCATGACATCACTTACACAAGCAAGTGGCTTAAT	120
Qy	138	GCACCGCTTGAAGCTCTCTCTCCCGCTGCTGTGCTGCGGGGAGCATCTTACAAAGCTG	197

Db	121	GCCTGCTTAGGCTCTCTTCTCCCGCTGGTGTGTCTGGCGGGGAGCATCTCAAGACTG	180
QY	198	CTATATGGCCAGTTCTTATCTTCTCCGTCTGTCTGTACTATCATCATCCGCTTTATTATAG	257
Db	181	CTATATGGCGAGTTCTCATCTTCTCCGTCTGTCTGTACTCATCATCCGCTTTATTATAG	240
QY	258	CTGGCCCTCAGCGAAGAACACAGCTGATGTGTTGAAACACTCTGTATTGCGACAGC	317
Db	241	CTGGCCCTCAGCGAAGAACACAGCTGATGTGTTGAAACACTCTGTATTGCGACAGC	300
QY	318	TACATCCAGCTCATCCCATTTCTTCTGTGTGGGCTTCTACGTACCGCTGTCTGACC	377
Db	301	TACATCCAGCTCATCCCATTTCTTCTGTGTGGGCTTCTACGTACCGCTGTCTGACC	360
QY	378	CGCGGTGTGAACCAAGTACGAGAACCTGGCGGTGGCCGACCGGCTATGAGCGCTGGGTG	437
Db	361	CGCGGTGTGAACCAAGTACGAGAACCTGGCGGTGGCCGACCGGCTATGAGCGCTGGGTG	420
QY	438	GGCTTCTGTGAAGCAAGAGACAGACAGGCGGCGTCTGTGGCGGACGCTCATCCGCTAC	497
Db	421	GGCTTCTGTGAAGCAAGAGACAGACAGGCGGCGTCTGTGGCGGACGCTCATCCGCTAC	480
QY	498	GCCCAACTGG 507	
Db	481	GCCCAACTGG 490	

RESULT 15			
BM685122/c			
LOCUS	BM685122	503 bp	linear
DEFINITION	UT-E-EJ1 a.j.1-09-0-UT.s1 UT-E-EJ1 Homo sapiens cDNA clone		
	UT-E-EJ1 a.j.1-09-0-UT 3', mRNA sequence.		

ACCESSION	BM685122	GI:18995018
VERSION	BM685122.1	
KEYWORDS	EST.	

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
Homo sapiens	1 (bases 1 to 503)	Bonaldo,M.F., Lennon,G. and Soares,M.B.	Normalization and subtraction: two approaches to facilitate gene discovery	Genome Res. 6 (9), 791-806 (1996)	9704447	Contact: Soares, MB
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.						

Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mpoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA-yes.

```

FEATURES
source
location/Qualifiers
1..503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-a)-1-09-0-UI"
/clone_1lb="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)
/note-Organ: eye;Vector: pT73-Pac (Pharmacia) with a

```

modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ1 is a subcloned cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)₁₈ tail. The
 sequence tags for this library are: fetal eyes, AGAATCAGA
 ; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;
 optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
 Macular, GTCC; RPE and Choroid, ACCCTA. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI).
 TAG_LIB-UI-E-EJ1
 TAG_TISSUE-RPE and Choroid
 TAG_SEO-ACCCTA"

BASE COUNT 96 a 107 c 123 g 177 t
 ORIGIN

Query Match 21.3%; Score 475.4; DB 14; Length 503;
 Best Local Similarity 99.8%; Pred. No. 9,7e-70;
 Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1547 AGAACCATCAGCGCCGTCAGAAAGCTTACAGTGTACAGGCGATGACACCAAGACAAAG 1606
 |||||||
 Db 494 AGAACCATCAGCGCCGTCAGAAAGCTTACAGTGTACAGGCGATGACACCAAGACAAAG 435
 |||||||
 QY 1607 CTTAAAGCTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGG 1666
 |||||||
 Db 434 CTTAAAGCTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGG 375
 |||||||
 QY 1667 GGCCTTGATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACCTTGAGAGTTAACT 1726
 |||||||
 Db 374 GGCCTTGATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACCTTGAGAGTTAACT 315
 |||||||
 QY 1727 GACGGATATGCCAGAGATCCCGAAAGTACCTCAAGAACTTTGAAACAATCACCAAC 1786
 |||||||
 Db 314 GACGGATATGCCAGAGATCCCGAAAGTACCTCAAGAACTTTGAAACAATCACCAAC 255
 |||||||
 QY 1787 CAACATACACACTACATCAAGAAATCATATGATGATCTTATGGGCTTGAAACACAGGA 1846
 |||||||
 Db 254 CAACATACACACTACATCAAGAAATCATATGATGATCTTATGGGCTTGAAACACAGGA 195
 |||||||
 QY 1847 TGAAGCACATTCCTAACCTGCTTCTATGGGATGCTTCCGACGACGAGTCTCACCCTG 1906
 |||||||
 Db 194 TGAAGCACATTCCTAACCTGCTTCTATGGGATGCTTCCGACGACGAGTCTCACCCTG 135
 |||||||
 QY 1907 TGTGTACACAGCAGAGACTGATCCAGTCAAGCCATACAGCTCTCCACACTGAAGAAC 1966
 |||||||
 Db 134 TGTGTACACAGCAGAGACTGATCCAGTCAAGCCATACAGCTCTCCACACTGAAGAAC 75
 |||||||
 QY 1967 GTGCCCTACACAGCAGCTGAATCAATGCTTAAATAGATAAATAATCCAGACTA 2023
 |||||||
 Db 74 ATGTCTTACACAGCAGCTGAATCAATGCTTAAATAGATAAATAATCCAGACTA 18

Search completed: June 9, 2003, 03:30:02
 Job time : 2956 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:30:48 ; Search time 5619.81 Seconds

(without alignments)
11543.117 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 cagggagtcacccacagccta.....aaaaaaaaaaaaaaaaaaaaa 2229

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: GenBank1:

1: gb_da:*
2: gb_da:*
3: gb_da:*
4: gb_da:*
5: gb_da:*
6: gb_da:*
7: gb_da:*
8: gb_da:*
9: gb_da:*
10: gb_da:*
11: gb_da:*
12: gb_da:*
13: gb_da:*
14: gb_da:*
15: gb_da:*
16: gb_da:*
17: gb_da:*
18: gb_da:*
19: gb_da:*
20: gb_da:*
21: gb_da:*
22: gb_da:*
23: gb_da:*
24: gb_da:*
25: gb_da:*
26: gb_da:*
27: gb_da:*
28: gb_da:*
29: gb_da:*
30: gb_da:*
31: gb_da:*
32: gb_da:*
33: gb_da:*
34: gb_da:*
35: gb_da:*
36: gb_da:*
37: gb_da:*
38: gb_da:*
39: gb_da:*
40: gb_da:*
41: gb_da:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208.4	99.1	2210	AF057169	AF057169 Homo sapi
2	2166.2	97.2	2171	AF073501	AF073501 Homo sapi
3	2002.4	89.8	2420	AF057170	AF057170 Homo sapi
4	861.4	38.6	1289	AY064707	AY064707 Sus scrofa
5	642.2	28.8	142092	AF139813	AF139813 Homo sapi
6	642.2	28.8	163915	AC087451	AC087451 Homo sapi
7	642.2	28.8	166867	AP003733	AP003733 Homo sapi
8	642.2	28.8	196080	AC004228	AC004228 Homo sapi
9	640.6	28.7	706	HSVMD2P10	HSVMD2P10 Homo sapi
10	640.6	28.7	112309	AC003025	AC003025 Homo sapi
11	640.6	28.7	133683	AC008487	AC008487 Homo sapi
12	576.8	25.9	1908	AF440756	AF440756 Homo sapi
13	552.8	24.8	160169	AC051664	AC051664 Homo sapi
14	538.4	24.2	1956	BC031186	BC031186 Homo sapi
15	538.4	24.2	1957	BC019528	BC019528 Homo sapi
16	535	24.0	1861	AK000139	AK000139 Homo sapi
17	459.2	20.6	2500	AK096459	AK096459 Homo sapi
18	457.6	20.5	1506	AF440757	AF440757 Homo sapi
19	446.2	20.0	2045	AF440757	AF440757 Homo sapi
20	358.4	16.1	492	HSVMD2P11	HSVMD2P11 Homo sapi
21	358.4	16.1	106648	AC004588	AC004588 Homo sapi
22	314.8	14.1	2874	AF218817	AF218817 Homo sapi
23	314.8	14.1	3610	AY061546	AY061546 Homo sapi
24	240.8	10.8	413	HSVMD2P04	HSVMD2P04 Homo sapi
25	240.8	10.8	1666	AF052095	AF052095 Homo sapi
26	218.6	9.8	32022	AC019521	AC019521 Homo sapi
27	218.6	9.8	170675	AC008139	AC008139 Homo sapi
28	218.6	9.8	192055	AC009183	AC009183 Homo sapi
29	218.6	9.8	221888	AE003686	AE003686 Homo sapi
30	191.6	8.6	246	AF057171	AF057171 Mus muscu
31	189	8.5	160169	HSVMD2P02	HSVMD2P02 Homo sapi
32	179.4	8.0	125804	AC051664	AC051664 Homo sapi
33	179.4	8.0	163514	AC019853	AC019853 Homo sapi
34	179.4	8.0	274351	AE003531	AE003531 Homo sapi
35	179.4	8.0	274351	AE003531	AE003531 Homo sapi
36	175	7.9	438	G05922	G05922 human STS W
37	175	7.9	1119	H0MEER2H	H0MEER2H human ferri
38	175	7.9	1198	AX409527	AX409527 Sequence
39	175	7.9	1198	H0MEER2H	H0MEER2H human ferri
40	161.4	7.2	239	HSVMD2P05	HSVMD2P05 Homo sapi
41	160.4	7.2	295	HSVMD2P07	HSVMD2P07 Homo sapi
42	155	7.0	289	HSVMD2P09	HSVMD2P09 Homo sapi
43	139.8	6.3	125804	AC019853	AC019853 Homo sapi
44	139.8	6.3	163514	AC091227	AC091227 Homo sapi
45	139.8	6.3	274351	AE003531	AE003531 Homo sapi

ALIGNMENTS

RESULT 1
AF057169
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
complete cds.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 (bases 1 to 2210)
Petrukhin, K., Kolst, M.J., Bakall, B., Li, W., Xie, G., Marknell, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,

Bergen, A.A., McGarity-Dugan, V., Figueroe, D., Austin, C.P., Metzker, M.L., Caskey, C.T. and Madellus, C.
 Identification of the gene responsible for Best macular dystrophy
 Nat. Genet. 19 (3), 241-247 (1998)
 MEDLINE 96324772
 PUBMED 9662395
 REFERENCE 2 (bases 1 to 2210)
 AUTHOR S Petrukhin, K.
 TITLE Direct Submission
 JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA
 FEATURES
 source Location/Qualifiers
 1. 2210
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1. 2210
 /gene="VMD2"
 105. 1862
 /note="Best's macular dystrophy gene; BMD; alternatively splice isoform"
 /product="bestrophin"
 /protein_id="AAC64343.1"
 /db_xref="GI:3335159"
 /translation="MTITYSQVNAIRITLQYDSTYQILPISFVLEVTYLVVTRWNNQYENLPWPRMSLVSGFVDEQGRILRLIRLANLNLILKSTAVYKRPSPQHLVQACPTFAHEKLEKLSLPHNMFVPMWVWNLNLSKMLGRIIDPLDLSLNMETLRTOGHLVAVDWISIPLYVOVTVAVYSFELCLVGRQFLNPAKAYPGHEIDLVAVVEFLQFFYVGMKVAEQLINPGEEDDDFEINMIVDRNLQVSLAVDEHODLPMPEDMYNKKPEPOPYTAASQFRASFGSTFNISLKEKEERQPNODEEDDAAGITGRFLGQSHDHHPRANSRTKLMPKRESILHGLPKNHRKAQNVNGQEDNKKAKLAVDAFKAPLYORGYVSAPOTPLSPMPFPLEPSPSKIHSTGIDTDKSKIVSSGAKSFELSESQALMEHPVSVQNRKIVFENLTMPELNHLEPLEQSPNTVHTLKDHPDPTWALENREDAHS"

BASE COUNT 556 a 647 c 531 g 476 t
 ORIGIN

Query Match 99.1%; Score 2208.4; DB 9; Length 2210;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAGGAGTCCACACGCTAGTCCAGACCTTCTGTGGATCATCGACCACTGGAA 60
 1 CAGGAGTCCACACGCTAGTCCAGACCTTCTGTGGATCATCGACCACTGGAA 60
 61 CCCACCTGACCCAGCCACCTGCTGCAAGCCACTGCTGGCCATGACCATCTTACA 120
 61 CCCACCTGACCCAGCCACCTGCTGCAAGCCACTGCTGGCCATGACCATCTTACA 120
 121 CAAAGCAAGTGCATATGCCGCTTAGGCTCTTCCGGCTGCTGTGGTGGCGGG 180
 121 CAAAGCAAGTGCATATGCCGCTTAGGCTCTTCCGGCTGCTGTGGTGGCGGG 180
 121 CAAAGCAAGTGCATATGCCGCTTAGGCTCTTCCGGCTGCTGTGGTGGCGGG 180
 121 CAAAGCAAGTGCATATGCCGCTTAGGCTCTTCCGGCTGCTGTGGTGGCGGG 180
 181 GCAGCATCTACAAAGTGTCTATATGGCGAGTCTTAACTCTCTGCTGCTACTACATCA 240
 181 GCAGCATCTACAAAGTGTCTATATGGCGAGTCTTAACTCTCTGCTGCTACTACATCA 240
 181 GCAGCATCTACAAAGTGTCTATATGGCGAGTCTTAACTCTCTGCTGCTACTACATCA 240
 241 TCCGCTTATTTATAGGCTGGCCCTCAGCGAAGAAACAACAGCTGATGTTAGAAACTGA 300
 241 TCCGCTTATTTATAGGCTGGCCCTCAGCGAAGAAACAACAGCTGATGTTAGAAACTGA 300
 301 CTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTGCTGTGGCTTTCTACG 360
 301 CTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTGCTGTGGCTTTCTACG 360
 301 CTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTGCTGTGGCTTTCTACG 360
 361 TGAGGCTGTGCTGACCCGCTGTGGTGAACAGTACGAGAACCTGCGGTGGCCGACCGCC 420
 361 TGAGGCTGTGCTGACCCGCTGTGGTGAACAGTACGAGAACCTGCGGTGGCCGACCGCC 420

421 TCATGAGCTGTGTGGGCTTGTGTCGAGGCAAGGACGAGGACCGGCTGCTGGG 480
 421 TCATGAGCTGTGTGGGCTTGTGTCGAGGCAAGGACGAGGACCGGCTGCTGGG 480
 481 GCAGCTCATCCGCTACAGCCCAACCTGGGCAAGCTGCTCATCTGCGAGCTGACGACCG 540
 481 GCAGCTCATCCGCTACAGCCCAACCTGGGCAAGCTGCTCATCTGCGAGCTGACGACCG 540
 481 GCAGCTCATCCGCTACAGCCCAACCTGGGCAAGCTGCTCATCTGCGAGCTGACGACCG 540
 541 CAGTCTACAAAGGCTTCCGACGCGGACCTGGTGTGAAGAGGCTTTATGACTCCGG 600
 541 CAGTCTACAAAGGCTTCCGACGCGGACCTGGTGTGAAGAGGCTTTATGACTCCGG 600
 541 CAGTCTACAAAGGCTTCCGACGCGGACCTGGTGTGAAGAGGCTTTATGACTCCGG 600
 601 CAGAACACAAAGAGTGTGAAGAACTAGCTACACACAAATGTTCTGGGTGCTGGG 660
 601 CAGAACACAAAGAGTGTGAAGAACTAGCTACACACAAATGTTCTGGGTGCTGGG 660
 601 CAGAACACAAAGAGTGTGAAGAACTAGCTACACACAAATGTTCTGGGTGCTGGG 660
 661 TGTGTTTCCCAACCTGTCATTAAGAGGCTGGCTGGAGTGGAAATCCGGACCTATCC 720
 661 TGTGTTTCCCAACCTGTCATTAAGAGGCTGGCTGGAGTGGAAATCCGGACCTATCC 720
 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACCTTGGCTAGTGTGAGACCTGTATG 780
 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACCTTGGCTAGTGTGAGACCTGTATG 780
 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACCTTGGCTAGTGTGAGACCTGTATG 780
 781 CCTACGACTGGATTAGTATCCCATCTGTGTATACACAGTGTGACTGTGGGCTGTACA 840
 781 CCTACGACTGGATTAGTATCCCATCTGTGTATACACAGTGTGACTGTGGGCTGTACA 840
 781 CCTACGACTGGATTAGTATCCCATCTGTGTATACACAGTGTGACTGTGGGCTGTACA 840
 841 GCTTCTTCCTGACTGTCTAGTGTGGGCGGCAATTTCTGAACCCAGCAAGGCTACCCG 900
 841 GCTTCTTCCTGACTGTCTAGTGTGGGCGGCAATTTCTGAACCCAGCAAGGCTACCCG 900
 841 GCTTCTTCCTGACTGTCTAGTGTGGGCGGCAATTTCTGAACCCAGCAAGGCTACCCG 900
 901 GCATGAGCTGGACCTGCTGTGGCCGCTTACGCTTCTGCAATTTCTATATGTTG 960
 901 GCATGAGCTGGACCTGCTGTGGCCGCTTACGCTTCTGCAATTTCTATATGTTG 960
 901 GCATGAGCTGGACCTGCTGTGGCCGCTTACGCTTCTGCAATTTCTATATGTTG 960
 901 GCATGAGCTGGACCTGCTGTGGCCGCTTACGCTTCTGCAATTTCTATATGTTG 960
 961 GCTGCTGAAGTGCAGAGGAGCTCATCAACCCCTTGGAGAGATGATGATTTTG 1020
 961 GCTGCTGAAGTGCAGAGGAGCTCATCAACCCCTTGGAGAGATGATGATTTTG 1020
 961 GCTGCTGAAGTGCAGAGGAGCTCATCAACCCCTTGGAGAGATGATGATTTTG 1020
 1021 AGACCACTGATGTGTGACAGAGAAATTTGACAGGTGCTCCCTGTGGCTGTGATGATGC 1080
 1021 AGACCACTGATGTGTGACAGAGAAATTTGACAGGTGCTCCCTGTGGCTGTGATGATGC 1080
 1021 AGACCACTGATGTGTGACAGAGAAATTTGACAGGTGCTCCCTGTGGCTGTGATGATGC 1080
 1081 ACCAGGACCTCTCTGGATGAGCGCGGACATGTGGAATTAAGCCGACAGCCAGCC 1140
 1081 ACCAGGACCTCTCTGGATGAGCGCGGACATGTGGAATTAAGCCGACAGCCAGCC 1140
 1141 CCTACACAGTGTCTTCCGCGAGTTCGTCGAGCTCTTATGAGGCTCCACTTCAACA 1200
 1141 CCTACACAGTGTCTTCCGCGAGTTCGTCGAGCTCTTATGAGGCTCCACTTCAACA 1200
 1141 CCTACACAGTGTCTTCCGCGAGTTCGTCGAGCTCTTATGAGGCTCCACTTCAACA 1200
 1201 TCAGCTGTACAAAGAGAGATGAGATTCAGGCCAATCAGAGAGAGAGAGATGCTC 1260
 1201 TCAGCTGTACAAAGAGAGATGAGATTCAGGCCAATCAGAGAGAGAGAGATGCTC 1260
 1201 TCAGCTGTACAAAGAGAGATGAGATTCAGGCCAATCAGAGAGAGAGAGATGCTC 1260
 1261 ACCTGGCATATGTCGCGCTTCTAGGCTGAGCTCCAGTCCAGTACCATCTCCAGGG 1320
 1261 ACCTGGCATATGTCGCGCTTCTAGGCTGAGCTCCAGTCCAGTACCATCTCCAGGG 1320
 1261 ACCTGGCATATGTCGCGCTTCTAGGCTGAGCTCCAGTCCAGTACCATCTCCAGGG 1320
 1321 CAAACTCAAGAGCAAACTACTGTGCGCCAAAGAGGAAATCCCTTCTCCAGAGGCGCTGC 1380
 1321 CAAACTCAAGAGCAAACTACTGTGCGCCAAAGAGGAAATCCCTTCTCCAGAGGCGCTGC 1380
 1321 CAAACTCAAGAGCAAACTACTGTGCGCCAAAGAGGAAATCCCTTCTCCAGAGGCGCTGC 1380
 1381 CCAAAAACCAAGGAGCGCCAAACAGAACTGTAGGGCCAGAGAAACAAGGCTGTGA 1440
 1381 CCAAAAACCAAGGAGCGCCAAACAGAACTGTAGGGCCAGAGAAACAAGGCTGTGA 1440
 1381 CCAAAAACCAAGGAGCGCCAAACAGAACTGTAGGGCCAGAGAAACAAGGCTGTGA 1440
 1441 AGCTTAAGGCTGTGAGGCTTCAAGTCTGAGCCAGTATCAGAGGCGAGGCTACTACA 1500
 1441 AGCTTAAGGCTGTGAGGCTTCAAGTCTGAGCCAGTATCAGAGGCGAGGCTACTACA 1500
 1441 AGCTTAAGGCTGTGAGGCTTCAAGTCTGAGCCAGTATCAGAGGCGAGGCTACTACA 1500
 1501 GTGCCCCACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACCATCAGCGC 1560

```

Db      1501 GTGCCCCAGACAGCCCCCTCAGCCCCCTCCAGCTTCTTCCCTTGAGACCATCAGCGC 1560
Oy      1561 CGTCAAGGCTTCACAGTGTACAGGATAGACACCAAGAACAAAGCTTAAAGACTGTGA 1620
Db      1561 CGTCAAGGCTTCACAGTGTACAGGATAGACACCAAGAACAAAGCTTAAAGCTGTGA 1620
Oy      1621 GTTCTGGGGCCCAAGAAAGTTTGAANTGCTTCAGAGACGATGGGCTTGTATGAGC 1680
Db      1621 GTTCTGGGGCCCAAGAAAGTTTGAANTGCTTCAGAGACGATGGGCTTGTATGAGC 1680
Oy      1681 ACCCAAGATATCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCCAG 1740
Db      1681 ACCCAAGATATCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCCAG 1740
Oy      1741 AGATCCCCGAAATTCACCTCAAGAAACCTTTGACACATCCACCAACATACACACTA 1800
Db      1741 AGATCCCCGAAATTCACCTCAAGAAACCTTTGACACATCCACCAACATACACACTA 1800
Oy      1801 CACTCAAGATACATGATGATCTTATTTGGGCTTTGAAAACAGGATGAAGCATTCCT 1860
Db      1801 CACTCAAGATACATGATGATCTTATTTGGGCTTTGAAAACAGGATGAAGCATTCCT 1860
Oy      1861 AACCTGCTCTTAAGGGGATGCTTCGCCAGCCAGGCTCTCACTGTGTACACACAGA 1920
Db      1861 AACCTGCTCTTAAGGGGATGCTTCGCCAGCCAGGCTCTCACTGTGTACACACAGA 1920
Oy      1921 GGACACTGATCCAGTACAGCATACAGCTGTCACACACAGAAAGCTGTCTTACACAG 1980
Db      1921 GGACACTGATCCAGTACAGCATACAGCTGTCACACACAGAAAGCTGTCTTACACAG 1980
Oy      1981 CCTGAATCAATGGTTAGCTTAATAGATAAATCCAGACTACTTACGCTTATATGCC 2040
Db      1981 CCTGAATCAATGGTTAGCTTAATAGATAAATCCAGACTACTTACGCTTATATGCC 2040
Oy      2041 TTTTATATATAAACTGTGAAGCTAGACTGACACCAATGGAATTAATTAAGTACAGTC 2100
Db      2041 TTTTATATATAAACTGTGAAGCTAGACTGACACCAATGGAATTAATTAAGTACAGTC 2100
Oy      2101 TGGATTCAGAGTCGGGAACCTTATAGTCTATCTGAAATCAAGACGCCACACTATAT 2160
Db      2101 TGGATTCAGAGTCGGGAACCTTATAGTCTATCTGAAATCAAGACGCCACACTATAT 2160
Oy      2161 ACTGCCCAACTAATGAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2210
Db      2161 ACTGCCCAACTAATGAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2210

RESULT 2
AF073501      2171 bp      mRNA      linear      PRI 02-SEP-1998
LOCUS      Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
DEFINITION      complete cds.
ACCESSION      AF073501
VERSION      AF073501.1 GI:3511241
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2171)
AUTHORS      Stohr, H., Marguardt, A., Rivera, A., Cooper, P. R., Nowak, N. J.,
TITLES      Shows, T. B., Gerhard, D. S., and Weber, B. H.,
A gene map of the Best's vitelliform macular dystrophy region in
JOURNAL      Chromosome Res. 8 (1), 48-56 (1998)
MEDLINE      98112782
PUBMED      9445487
REFERENCE      2 (bases 1 to 2171)
AUTHORS      Marguardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A., and
TITLES      Weber, B. H. F.,
Direct Submission
JOURNAL      Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am

```

```

FEATURES
Hubland, Wuerzburg 97074, Germany
Location/Qualifiers
1..2171
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/tissue-type="retinal pigment epithelium"
1..2171
/gene="VMD2"
89..1846
/gene="VMD2"
/codon_start=1
/product="vitelliform macular dystrophy protein"
/db_xref="GI:3511242"
/translation="MTITYSQVANNRLGFSRLLCMRGSYIKLLYGEFLFLCY
IIRFYRLALTEEOQIMEKTLICDSTYIOLIPISFLGFTYLVYTHMNOYENLPM
PRLMSIVSGFEVGEDEOGRLLRLTIRVANGNYLILRSVTAAYKRPSAOHLVOA
GEMTFAEHKOLEKLSLPHMFEVPPVWPVFPANLSMKALGSRINDPILLOSLEMMTLR
TQCGHLYAYDWTISILVYNOYTVAYVSFFLTCVGRPLNPKATPGHELDLVPEF
TELOREFYVGMKVAIBOLINPRGEDDDEETNMIVDRNLQVSLAVDEMDHDLPRMEP
DMYWKPEQPPYTAASQFRASFWGSPFNLSLKEEMEPQNOEDEDAGLIGR
FLGDSHDIHHPRANRKLMPKRESLHLELPHNKAANKONVROEDNKAAMKLKAV
DAFKSAPLYORPGYSAPOTPLSPMPFPLPSPSKLSHVSYGLDTRKSLKTVSSG
AKKSFELSESQDGLMEHPYEQVARKTYEFLVLMPELPEHNLKEPLEDSPTNHTT
LKDHPDPAWALENREDAHS"
BASE COUNT      540 a      638 c      523 g      470 t
ORIGIN
Query Match      97.2%; Score 2166.2; DB 9; Length 2171;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy      17 CCTATGCCGACACTTCTGTGGGATCATCGGACCCACTGGAAACCCACTGACCCAAG 76
Db      1 CCTATGCCGACACTTCTGTGGGATCATCGGACCCACTGGAAACCCACTGACCCAAG 60
Oy      77 CCCACCTGCTCAGGCCACTGCTGGGACATGACATCAATCAACCAAGTGGCTTA 136
Db      61 CCCACCTGCTCAGGCCACTGCTGGGACATGACATCAATCAACCAAGTGGCTTA 120
Oy      137 TGCCTGCTTAAAGCTCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 196
Db      121 TGCCTGCTTAAAGCTCTCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Oy      197 GCTATAGGGAGGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
Db      181 GCTATATGGGAGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Oy      257 GCTGCGCTCAGGAGAACACAGCTGATGTTGAGAAATGACTGTTGCGACAG 316
Db      241 GCTGCGCTCAGGAGAACACAGCTGATGTTGAGAAATGACTGTTGCGACAG 300
Oy      317 CTACATCCAGCTATCCCAATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
Db      301 CTACATCCAGCTATCCCAATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Oy      377 CCGCTGTGAGACAGAGAGAAACCTGCGCTGCGGACCGGCTCATGAGCTGTGTC 436
Db      361 CCGCTGTGAGACAGAGAGAAACCTGCGCTGCGGACCGGCTCATGAGCTGTGTC 420
Oy      437 GGGCTTCGTGGAAGGACAGAGAGCGCGCTGCTGCGGACCGGCTCATGAGCTGTG 496
Db      421 GGGCTTCGTGGAAGGACAGAGAGCGCGCTGCTGCGGACCGGCTCATGAGCTGTG 480
Oy      497 CGCACAACCTGGGAGAAAGTGTCTATCTGCGACGCTGACACCGGAGTTCACAGCGCTT 556
Db      481 CGCACAACCTGGGAGAAAGTGTCTATCTGCGACGCTGACACCGGAGTTCACAGCGCTT 540
Oy      557 CCCACGCGCCAGACACTGTGTGACAGGCTTATGACTCCGCGAGAACACAGAGATT 616

```

Db 541 CCCACGCCCCAGACCTGCTGTCAGACGAGCTTATGACTCCGGGCAAGACAGACT 600
 Oy 617 GGAGAACTGAGCCCTACACACACATGTTCTGGGGTCCCTGGGTGGTTGGCCAACT 676
 Db 601 GGAGAACTGAGCCCTACACACACATGTTCTGGGGTCCCTGGGTGGTTGGCCAACT 660
 Oy 677 GTCAATGAAAGCGGTGGAGTGCATCCGGGACCTATCTGCTCCAGAGCGCTGT 736
 Db 661 GTCAATGAAAGCGGTGGAGTGCATCCGGGACCTATCTGCTCCAGAGCGCTGT 720
 Oy 737 GAACAGATGAACACCTTGCCTACAGTGTGAGACACCTGTATGCTTACAGTGA 796
 Db 721 GAACAGATGAACACCTTGCCTACAGTGTGAGACACCTGTATGCTTACAGTGA 780
 Oy 797 TATCCACGTGGTGTATACAGAGTGTGATCTGGCGGTGATACCTTCTTCTGACT 856
 Db 781 TATCCACGTGGTGTATACAGAGTGTGATCTGGCGGTGATACCTTCTTCTGACT 840
 Oy 857 TCTAGTTGGGGGAGTTCTGMAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGCT 916
 Db 841 TCTAGTTGGGGGAGTTCTGMAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGCT 900
 Oy 917 CGTGTGCCCTCTTACAGTGTCTGACCTTCTTCTTATGTTGGCTGGCTGAAGTGGC 976
 Db 901 CGTGTGCCCTCTTACAGTGTCTGACCTTCTTCTTATGTTGGCTGGCTGAAGTGGC 960
 Oy 977 AGAGAGCTCATCAACCCCTTGGAGAGATGATATTTTGGAGCACTGATGTGT 1036
 Db 961 AGAGAGCTCATCAACCCCTTGGAGAGATGATATTTTGGAGCACTGATGTGT 1020
 Oy 1037 CGACAGAAATTTGCAAGTGTCTGCTGCTGCTGCTGATGATGACACAGACCTGCTCG 1096
 Db 1021 CGACAGAAATTTGCAAGTGTCTGCTGCTGCTGCTGATGATGACACAGACCTGCTCG 1080
 Oy 1097 GATGAGCGCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
 Db 1081 GATGAGCGCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 Oy 1157 CGCCAGTTCCGTGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGMAACAAGA 1216
 Db 1141 CGCCAGTTCCGTGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGMAACAAGA 1200
 Oy 1217 GGAATGAGATTTCCAGCCCAATCAGAGAGAGAGAGAGATGCTCAAGCTGATATTGG 1276
 Db 1201 GGAATGAGATTTCCAGCCCAATCAGAGAGAGAGAGAGATGCTCAAGCTGATATTGG 1260
 Oy 1277 CGGCTTCCAGGCGTGGAGCTCCATGATCCATCCTCCAGGCGCAAACTAAGGACCA 1336
 Db 1261 CGGCTTCCAGGCGTGGAGCTCCATGATCCATCCTCCAGGCGCAAACTAAGGACCA 1320
 Oy 1337 ACTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAG 1396
 Db 1321 ACTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Oy 1397 AGCCAAACAGAACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1456
 Db 1381 AGCCAAACAGAACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Oy 1457 CGGCTTCCAGGCGTGGAGCTCCATGATCCATCCTCCAGGCGCAAACTAAGGACCA 1516
 Db 1441 CGGCTTCCAGGCGTGGAGCTCCATGATCCATCCTCCAGGCGCAAACTAAGGACCA 1500
 Oy 1517 CGTACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1576
 Db 1501 CGTACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Oy 1577 TGTACAGGAGATGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1636
 Db 1561 TGTACAGGAGATGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Oy 1637 AAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1696
 Db 1621 AAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680

Oy 1697 AGTGAAGAGAAAGCTGTGAGTTAACTGACGATATATCCAGAGATCCCGAAATCA 1756
 Db 1681 AGTGAAGAGAAAGCTGTGAGTTAACTGACGATATATCCAGAGATCCCGAAATCA 1740
 Oy 1757 CCTCAAGAAAGCTTGGAAACATCAACCAACCAATACACTACACTACAAAGATCACT 1816
 Db 1741 CCTCAAGAAAGCTTGGAAACATCAACCAACCAATACACTACACTACAAAGATCACT 1800
 Oy 1817 GGATCCTTATTTGGGCTTGGAAACAGGAGATGAACACATTCCTTAACCTGCTTCTATG 1876
 Db 1801 GGATCCTTATTTGGGCTTGGAAACAGGAGATGAACACATTCCTTAACCTGCTTCTATG 1860
 Oy 1877 GGGATGCTTGGCCAGCCAGGCTCTCACTGTGTGATACACAGACAGACATGATCAGTC 1936
 Db 1861 GGGATGCTTGGCCAGCCAGGCTCTCACTGTGTGATACACAGACAGACATGATCAGTC 1920
 Oy 1937 ACAGCCATACAGCTGTCCACACTGAAAGACGTGTCTTACACAGCTGAAATGATGTT 1996
 Db 1921 ACAGCCATACAGCTGTCCACACTGAAAGACGTGTCTTACACAGCTGAAATGATGTT 1980
 Oy 1997 AGCTTAATAGATTAATAATCCAGACTACTACTGAGCTTAAATGCTTTATTCATTAATAAC 2056
 Db 1981 AGCTTAATAGATTAATAATCCAGACTACTACTGAGCTTAAATGCTTTATTCATTAATAAC 2040
 Oy 2057 TGTGAAGCTGACTGAACCATGGAACATTTAATCACTAGACTGTGATTCAGAGTCGGG 2116
 Db 2041 TGTGAAGCTGACTGAACCATGGAACATTTAATCACTAGACTGTGATTCAGAGTCGGG 2100
 Oy 2117 AACCTTATGCTTATCTGAATCAAGACAGACAGCTTATGATATGCTGCCAACTAATG 2176
 Db 2101 AACCTTATGCTTATCTGAATCAAGACAGACAGCTTATGATATGCTGCCAACTAATG 2160
 Oy 2177 AGTTAATATA 2187
 Db 2161 AGTTAATATA 2171

RESULT 3
 AF057170
 LOCUS
 DEFINITION Homo sapiens bestraphin (VMD2) mRNA, alternatively spliced product,
 complete cds.
 ACCESSION AF057170
 VERSION AF057170.1 GI:3335160
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2420)
 Petrunkhin, K., Koist, M. J., Bakall, B., Li, W., Xie, G., Marknell, T.,
 Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vuojic, M.,
 Bergen, A. A., McGarry-Dugan, V., Figueroa, D., Austin, C. P.,
 Metzker, M. L., Caskey, C. T. and Madellus, C.
 Identification of the gene responsible for Best macular dystrophy
 Nat. Genet. 19 (3), 241-247 (1998)
 JOURNAL MEDLINE
 PUBMED 96324772
 9662395
 2 (bases 1 to 2420)
 REFERENCE
 AUTHOR Petrunkhin, K.
 TITLE Direct Submission
 JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
 Laboratories, West Point, PA 19486, USA
 FEATURES
 source
 1..2420
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1..2420
 /gene="VMD2"
 105..1412
 CDS

		/gene="VMD2" /note="Best's macular dystrophy gene; BMD; alternatively splice isoform; contains exon 7" /codon_start=1 /product="bestrophin" /protein_id="AAC64344.1" /db_xref="gi:3335161" /translation="MTITVTSOVAMNARLSSPRLLCWRGSIYKLYLGEPLIFLCY IIRPIYALTEEOQLMEKTLTXCDYIOLIPISFVLTIVLVTRMNOENLFW PRLMSLVSGFKDQRLKRLTIRIVNLIRSVTAIVYRFPASQHLVQA GEMTPRHKOLEKLSPHNFVPMWFANLSKAMIGRIIDPILLOSILNENITR TQCGHLEIDMISIPLYTQVYVAVYSEFLCLVGHQFLNPAKAYGHEDLVVPR TQLEFFVGMLEKVLGRHGRGQOULETRMOCERKYSRVSSQAMWPT PVIPATREAVAGESELEGRRLMMOSSSLPRLMMMLILPTGICICPCFLMWR CRRCLGMSRSTCISPSHSPROLLPSSVPEPLVNAFPSTSA"	
BASE COUNT	605 a 695 c 612 g 508 t		
ORIGIN			
Query Match	89.8%; Score 2002.4; DB 9; Length 2420;		
Best Local Similarity	91.6%; Pred. No. 0;		
Matches 2216; Conservative	0; Mismatches 1; Indels 203; Gaps 1;		
OY	1	CAGGGAGTCCCAACACCTAGTGCAGACCTTCTGGGATCATGGACCACTGGAA	60
Db	1	CAGGGAGTCCCAACACCTAGTGCAGACCTTCTGGGATCATGGACCACTGGAA	60
OY	61	CCCCACCTGACCCCAACCCACCTGCTGACGCCACCTGGGCATGACATCATTTACA	120
Db	61	CCCCACCTGACCCCAACCCACCTGCTGACGCCACCTGGGCATGACATCATTTACA	120
OY	121	CAGGCCAGTGGCTAATGCGCCGCTTAAAGCTCTTCCGCCCTGCTGCTGCTGGGG	180
Db	121	CAGGCCAGTGGCTAATGCGCCGCTTAAAGCTCTTCCGCCCTGCTGCTGCTGGGG	180
OY	181	GCAGCATCTACAGCTGCTATATAGGAGTCTTATCTCTGCTGCTCTCTACTATCA	240
Db	181	GCAGCATCTACAGCTGCTATATAGGAGTCTTATCTCTGCTGCTCTCTACTATCA	240
OY	241	TCCGCTTATTTATATAGGCTGGCCCTACAGGAACAACAGCTGATGTTTGAAGAAC	300
Db	241	TCCGCTTATTTATATAGGCTGGCCCTACAGGAACAACAGCTGATGTTTGAAGAAC	300
OY	301	CTCTGATTTGGAGACGTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTAC	360
Db	301	CTCTGATTTGGAGACGTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTAC	360
OY	361	TGAGCGTGGTGTGACCCGCTGCTGAGAACCAAGTACAGAACTGCGCGTGGCCGCG	420
Db	361	TGAGCGTGGTGTGACCCGCTGCTGAGAACCAAGTACAGAACTGCGCGTGGCCGCG	420
OY	421	TGATGAGCGCTGGTGGGCTTCTGAGAGCAAGAGCAGCAAGCGCGCTGCGGGC	480
Db	421	TGATGAGCGCTGGTGGGCTTCTGAGAGCAAGAGCAGCAAGCGCGCTGCGGGC	480
OY	481	GCAGCGTCATCGGCTAGCCCAACCTGGGCAACGTCTCATCTGCGAGCGTCAAGCA	540
Db	481	GCAGCGTCATCGGCTAGCCCAACCTGGGCAACGTCTCATCTGCGAGCGTCAAGCA	540
OY	541	CAGTCTACAAAGCGTTCCCAAGCCCAAGCACTGGTGGCAAGAGGCTTTATGACTCG	600
Db	541	CAGTCTACAAAGCGTTCCCAAGCCCAAGCACTGGTGGCAAGAGGCTTTATGACTCG	600
OY	601	CAGAACCAAGAGCTGGAGAACTGAGCCTACCAACAACATTTCTGGGTGGCCCTGG	660
Db	601	CAGAACCAAGAGCTGGAGAACTGAGCCTACCAACAACATTTCTGGGTGGCCCTGG	660
OY	661	TGTGTTTGGCAACCTGTCAATGAAGCGTGGCTTGAAGTCAATCCGGGACCTATCC	720
Db	661	TGTGTTTGGCAACCTGTCAATGAAGCGTGGCTTGAAGTCAATCCGGGACCTATCC	720
OY	721	TGCTCAGAGACCTGCTGAGAGATGAAACACTTGGTACTCACTGAGTGGACACTGTAT	780
Db	721	TGCTCAGAGACCTGCTGAGAGATGAAACACTTGGTACTCACTGAGTGGACACTGTAT	780

OY	781	CCTAGACGTGATTAATATCCACTGGGNTATACACAGGTGTGACTGTGGCGGTAC	840
Db	781	CCTAGACGTGATTAATATCCACTGGGNTATACACAGGTGTGACTGTGGCGGTAC	840
OY	841	GCTTCTTCTGACTGTCTAGTGTGGGGGCGAGTTTCTGAACCCAGCCAGGCTTAC	900
Db	841	GCTTCTTCTGACTGTCTAGTGTGGGGGCGAGTTTCTGAACCCAGGCTTAC	900
OY	901	GCCATGAGCTGGAGACCTGTTGTGCCCTCTTACAGCTTCTGCACTTCTTATGTTG	960
Db	901	GCCATGAGCTGGAGACCTGTTGTGCCCTCTTACAGCTTCTGCACTTCTTATGTTG	960
OY	961	GCTGGCTGAA-----	970
Db	961	GCTGGCTGAAAGTGGGCTCTCCAGGGCCCTGCTGGCTGGAGCATGGCCAGAGGGTC	1020
OY	971	-----	970
Db	1021	ATGGCAGACAGTCTTGAAGACAGAGATGCAGTCTCAGAAAGAGTCTCAGCGTAG	1080
OY	971	-----	970
Db	1081	AAAGCAGCCAGGCTGTGTGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGGAG	1140
OY	971	-----	970
Db	1141	GAGAAATCGTTGAACCCGGGAGGCGGAGGTGTGTGTGGCAGAGACAGCTCATCA	1200
OY	998	TGAGAGGATGATGATTTTGAAGCAACCTGGAATGTGACAGAAATTTGAGAGTGC	1057
Db	1201	TGAGAGGATGATGATTTTGAAGCAACCTGGAATGTGACAGAAATTTGAGAGTGC	1260
OY	1058	CCGTGTGGCTGTGATGATGATGACAGAGACCGCTCGAGTGGAGACCCGACATGAC	1117
Db	1261	CCGTGTGGCTGTGATGATGATGACAGAGACCGCTCGAGTGGAGACCCGACATGAC	1320
OY	1118	GAAATAGCCCGAGCCAGCCAGCCCTTACACAGCTGCTTCCGCCAATTCCTGAGAC	1177
Db	1321	GAAATAGCCCGAGCCAGCCAGCCCTTACACAGCTGCTTCCGCCAATTCCTGAGAC	1380
OY	1178	CTTATGGGCTCCACTTCAATCATGAGCTGAAACAAAGAGATGAGATTCAGGCCAA	1237
Db	1381	CTTATGGGCTCCACTTCAATCATGAGCTGAAACAAAGAGATGAGATTCAGGCCAA	1440
OY	1238	TCAGAGAGACGAGAGATGCTCAGCTGGGCAATTTGGCGCTTCTGAGGCTGAGTC	1297
Db	1441	TCAGAGAGACGAGAGATGCTCAGCTGGGCAATTTGGCGCTTCTGAGGCTGAGTC	1500
OY	1298	CCATGATACCAATCTCCAGGGCAAACTCAAGAACCAAACTGATGAGCCCAAGAGGA	1357
Db	1501	CCATGATACCAATCTCCAGGGCAAACTCAAGAACCAAACTGATGAGCCCAAGAGGA	1560
OY	1358	ATCCCTTCTCAGAGAGGCGCTGCCAAAACCAAGCAGGCAACAGACGTTAGGGG	1417
Db	1561	ATCCCTTCTCAGAGAGGCGCTGCCAAAACCAAGCAGGCAACAGACGTTAGGGG	1620
OY	1418	CCAGAGACCAACAAAGGCTTGAAGCTTGAAGCTTGAAGGCTTCAAGTCTGGCGCT	1477
Db	1621	CCAGAGACCAACAAAGGCTTGAAGCTTGAAGCTTGAAGGCTTCAAGTCTGGCGCT	1680
OY	1478	GTAATCAGAGGCGAGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCTCATGTT	1537
Db	1681	GTAATCAGAGGCGAGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCTCATGTT	1740
OY	1538	CTTCCCTTGAACCATCAGAGCGCTGCAAAAGCTTCAAGTCTCAGAGCATGACACCA	1597
Db	1741	CTTCCCTTGAACCATCAGAGCGCTGCAAAAGCTTCAAGTCTCAGAGCATGACACCA	1800
OY	1598	AGACAAAGCTTAAAGACTGTGATTTGGGGCCCAAGAAAGTTTGAATGCTGAGA	1657
Db	1801	AGACAAAGCTTAAAGACTGTGATTTGGGGCCCAAGAAAGTTTGAATGCTGAGA	1860

Oy	1658	GAGCGATGGGCGCTTGATGGAGACCCAGGAAGTATCTCAAGTAGSAGAGAAGAAACTGTGGA	1711
Dd	1861	GAGCATGTGGGCGCTTGATGGAGACCACCGAAGTATCTCAAGTAGSAGAGAAGAAACTGTGGA	1920
Oy	1718	GTTTAACCTGACGGATATGCCAGAGATCCCAGAAAATCAGCTTCAAAAGAACTTTGGAAACA	1777
Dd	1921	GTTTTAACCTGACGGATATGCCAGAGATCCCCGAAAAATCACCTTCAAAAGAACTTTGGAAACA	1980
Oy	1778	ATCACCACCAACCAATACACACTACACTCAAAGATACATGATGATCTTATTGGGCGTTGGA	1837
Dd	1961	ATCACCACCAACCAATACACACTACACTCAAAGATACATGATGATCTTATTGGGCGTTGGA	2040
Oy	1838	AAACAGGGATGAAGCAGCATTCCTTAACCTGCTTCCATAATGGGAGATGCTCGCAGCCAGGT	1897
Dd	2041	AAACAGGGATGAAGCAGCATTCCTTAACCTGCTTCCATAATGGGAGATGCTCGCAGCCAGGT	2100
Oy	1898	CCTGACCTGTGTGTACACACAGCAGACACTATGCATGCACAGCCATACAGCTGTCCACA	1957
Dd	2101	CCTGACCTGTGTGTACACACAGCAGACACTATGCATGCACAGCCATACAGCTGTCCACA	2160
Oy	1958	CTGAAGAACGTGTCTCTACACAGCCTGATCAAAATGTTAGCTTAAATGATAAATATCC	2017
Dd	2161	CTGAAGAACGTGTCTCTACACAGCCTGATCAAAATGTTAGCTTAAATGATAAATATCC	2220
Oy	2018	AGACTACTTCAGCCTTTATGCGCTTTTATTCATATAAAACCTGTGAAGCTAGACTGAACCA	2077
Dd	2221	AGACTACTTCAGCCTTTATGCGCTTTTATTCATATAAAACCTGTGAAGCTAGACTGAACCA	2280
Oy	2078	TTGGAAACATTTAACTGACACTGTGGATTCGAGTCGAGCGGAAACCTTGATCTATCTGAAT	2137
Dd	2281	TTGGAAACATTTAACTGACACTGTGGATTCGAGTCGAGTCGAGCGGAAACCTTGATCTATCTGAAT	2340
Oy	2138	CCAAGACAGCCACACCTTAGTATACCTGCCAAACTAATGACTTTAATTAATACAAATACT	2197
Dd	2341	CCAAGACAGCCACACCTTAGTATACCTGCCAAACTAATGACTTTAATTAATACAAATACT	2400
Oy	2198	CGTTAAAAAATAAAAAAAAA 2217	
Dd	2401	CGTTAAAAAATAAAAAAAAA 2420	
RESULT 4			
.LOCUS	AY064707	1289 bp	mRNA linear MAM 19-AUG-2002
DEFINITION	Sus scrofa bestrophin mRNA, partial cds.		
ACCESSION	AY064707		
VERSION	AY064707.1	GI:18476495	
KEYWORDS	.		
SOURCE	Pig.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.		
TITLE	1 (bases 1 to 1289)		
JOURNAL	Marmorstein,L.Y., McLaughlin,P.J., Stanton,J.B., Yan,L., Crabbs,J.W.		
PubMed	and Marmorstein,A.D.		
AUTHORS	Bestrophin interacts Physically and Functionally with Protein Phosphatase 2A		
	J. Biol. Chem. 277 (34), 30591-30597 (2002)		
	12058047		
	2 (bases 1 to 1289)		
	Marmorstein,L.Y., McLaughlin,P.J., Stanton,B., Yan,L., Crabbs,J.W.		
	and Marmorstein,A.D.		
	Direct Submission		
	Submitted (30-NOV-2001) Cole Eye Institute, 131, Cleveland Clinic		
	Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA		
FEATURES	Location/Qualifiers		
source	1..1289		
	/organism="Sus scrofa"		
	/db_xref="taxon:9823"		
	<1..1289		
	/codon_start=3		
	/product="bestrophin"		
	/protein_id="AAI40862.1"		

Query Match	Best Local Similarity	Matches 1058:	Conservative	Score 861.4:	DB 4:	Length 1289:
				Pred. No. 1.6e-191:		
				0: Mismatches 221:	Indels 24:	Gaps 4:
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						
Gaps 4:						
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						
Gaps 4:						
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						
Gaps 4:						
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						
Gaps 4:						
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						
Gaps 4:						
BASE COUNT	293 a	390 c	360 g	246 t		
ORIGIN						
Query Match	38.6%;					
Best Local Similarity	81.2%;					
Matches 1058:						
Conservative						
Score 861.4:						
DB 4:						
Length 1289:						
Pred. No. 1.6e-191:						
0: Mismatches 221:						
Indels 24:						

[illegible]

	repeat_region	/rpl_family="SVA"	complement(47665)	.477773)	
	repeat_region	/rpl_family="Alu"	complement(47792)	.48077)	
	repeat_region	/rpl_family="Alu"	complement(48110)	.48388)	
	repeat_region	/rpl_family="Alu"	complement(48827)	.49102)	
	repeat_region	/rpl_family="SVA"	complement(49395)	.49626)	
	repeat_region	/rpl_family="Alu"	50184	.50437	
	repeat_region	/rpl_family="Alu"			
Query Match		28.8%	Score 642.2;	DB 9;	length 142092;
Best Local Similarity		99.5%	Pred.	No.1e-139;	
Matches 644;	Conservative	0;	Mismatches	3;	Indels 0; Gaps 0;
OY	1202	CAGCCTTAACAAAGAGAGATGAGATTCCAGGCCAATCAAGAGAGAGAGAGATGTCTCA	1261		
Dd	25481	CAGCCTTAACAAAGAGAGATGAGATTCCAGGCCAATCAAGAGAGAGAGAGATGTCTCA	25422		
OY	1362	CGCTGGCATTTGGCGCTTCTCTTAGGCCTGCAAGTCCTCATACACTCTCCAGGGC	1321		
Dd	25421	CGCTGGCATTTGGCGCTTCTCTTAGGCCTGCAAGTCCTCATACACTCTCCAGGGC	25362		
OY	1322	AAACTCAAGACCAAACTACTGTGGCCCAAGAAGGAATCCCTTCTCACAGAGGCTGCC	1381		
Dd	25361	AAACTCAAGACCAAACTACTGTGGCCCAAGAAGGAATCCCTTCTCACAGAGGCTGCC	25302		
OY	1382	CAAAAACACAAGAGAGCCCAAAGAAGCTTAGGGGCGAAGAACACAAAGAGGCTTGAA	1441		
Dd	25301	CAAAAACACAAGAGAGCCCAAAGAAGCTTAGGGGCGAAGAACACAAAGAGGCTTGAA	25242		
OY	1442	GCTTAAAGCTGTGGAGCCTTCAAGTCTGGGCCCACTGTATCAAGAGGCCAGGCTACTACG	1501		
Dd	25241	GCTTAAAGCTGTGGAGCCTTCAAGTCTGGGCCCACTGTATCAAGAGGCCAGGCTACTACG	25182		
OY	1502	TGCCCCACAGACGCCCTCTAGCCCACCTCCATGTTCTTCCCCCTGAACCATACAGCGCC	1561		
Dd	25181	TGCCCCACAGACGCCCTCTAGCCCACCTCCATGTTCTTCCCCCTGAACCATACAGCGCC	25122		
OY	1562	GTCAAAACCTTCACAGTGTCAAGAGCATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG	1621		
Dd	25121	GTCAAAACCTTCACAGTGTCAAGAGCATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG	25062		
OY	1622	TTCGGGGGCCAAGAAAAAGTTTTAAATTTGCTCTCAAGAGACCGATGGGGCCTTGATGAGCA	1681		
Dd	25061	TTCGGGGGCCAAGAAAAAGTTTTAAATTTGCTCTCAAGAGACCGATGGGGCCTTGATGAGCA	25002		
OY	1682	CCCAAGAGTATCTCAAGTGAAGAGAAAACTGTGGATTAACTTACAGCATATGCGCAGA	1741		
Dd	25001	CCCAAGAGTATCTCAAGTGAAGAGAAAACTGTGGATTAACTTACAGCATATGCGCAGA	24942		
OY	1742	GATCCCCGAAAATACCTCAAGAGACCTTTGGAACATACCAACCACATACACACTAC	1801		
Dd	24941	GATCCCCGAAAATACCTCAAGAGACCTTTGGAACATACCAACCACATACACACTAC	24882		
OY	1802	ACTCAAAAGATCACATGGATCCTTATTTGGGCTTGGAAAAACAGGATG 1848			
Dd	24881	ACTCAAAAGATCACATGGATCCTTATTTGGGCTTGGAAAAACAGGCTG 24835			
RESULT 6					
LOCUS	AC087451/c	163915 bp	DNA	linear	HMG 04-JUL-2001
DEFINITION	Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT				
ACCESSION	AC087451				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				

SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 163915)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 11, clone RP11-810P12
REFERENCE 2 (bases 1 to 163915)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Boukhvalter, B., Brown, A., Camarata, D., Campiano, A., Choquet, Y., Colangelo, M., Collins, S., Collumore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyne, S., Gande, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Maclean, C., Macdonald, P., Marquis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Reta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Teafaye, S., Theodore, J., Travers, M., Trevis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zaimon, J., Zambek, L., Zimmer, A. and Zody, M.
DIRECT SUBMISSION
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 4, 2001 this sequence version replaced g1:12039464.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996) 1997
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L11863
Center clone name: 810_P_12
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162458 bases at least Q40
Consensus quality: 162916 bases at least Q30
Consensus quality: 163173 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 163115; sum-of-ctrls
Quality coverage: 10.5 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 bases; sum-of-ctrls
NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 17285: contig of 17285 bp in length
* 17286 17385: gap of 100 bp
* 17386 19170: contig of 1785 bp in length
* 19171 19270: gap of 100 bp
* 19271 28600: contig of 9330 bp in length
* 28601 28700: gap of 100 bp
* 28701 47671: contig of 18971 bp in length
* 47672 47771: gap of 100 bp

FEATURES
source
* 47772 105063: contig of 57292 bp in length
* 105064 105163: gap of 100 bp
* 105164 142181: contig of 37018 bp in length
* 142182 142281: gap of 100 bp
* 142282 163915: contig of 21634 bp in length.
location/Qualifiers
1. 163915
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-810P12"
/clone_lib="RP11-810P12 Human Male BAC"
1. 17285
/note="assembly-fragment"
vector_end:sp6
vector_side:left"
17386..19170
/note="assembly-fragment"
19271..28600
/note="assembly-fragment"
28701..47671
/note="assembly-fragment"
47772..105063
/note="assembly-fragment"
105164..142181
/note="assembly-fragment"
142282..163915
/note="assembly-fragment"
clone_end:T7
vector_side:right"
BASE COUNT 43950 a 38090 c 37945 g 43329 t 601 others
ORIGIN
Query Match 28.8%; Score 642.2; DB 2; Length 163915;
Best Local Similarity 99.5%; Pred. No. 1e-139;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1202 CAGCCTAACAAGAGAGATGAGTTCACGCCCAATCAGAGAGAGAGATGATCA 1261
DB 151430 CAGCCTAACAAGAGAGATGAGTTCACGCCCAATCAGAGAGAGAGATGATCA 151371
QY 1262 CGCTGGCATATGCGCGCTTCTTCAAGCCCTGCACTCCCATGATCATTCCAGGCG 1321
DB 151370 CGCTGGCATATGCGCGCTTCTTCAAGCCCTGCACTCCCATGATCATTCCAGGCG 151311
QY 1322 AAACCTAAGAGACCAACTACTGTGCGCCAGAGAGATCCCTTCCACAGAGGCGCTGCC 1381
DB 151310 AAACCTAAGAGACCAACTACTGTGCGCCAGAGAGATCCCTTCCACAGAGGCGCTGCC 151251
QY 1382 CAAAAACCAAG 1441
DB 151250 CAAAAACCAAG 151191
QY 1442 GCTTAAGGCTGTGAGAGCCTTCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACTACAG 1501
DB 151190 GCTTAAGGCTGTGAGAGCCTTCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACTACAG 151131
QY 1502 TCCCCACAG 1561
DB 151130 TCCCCACAG 151071
QY 1562 GTCAAGGCTTCAAGTGTACAGAGATGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
DB 151070 GTCAAGGCTTCAAGTGTACAGAGATGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 151011
QY 1622 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 151010 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150951
QY 1682 CCCAGAACTATCTCAAGTGTGAG 1741
DB 150950 CCCAGAACTATCTCAAGTGTGAG 150891

OY 1742 GATCCCGAAATACCTCAAGAAACCTTTGGACATCACCACCAACATACACATACACATAC 1801
 DB 150890 GATCCCGAAATACCTCAAGAAACCTTTGGACATCACCACCAACATACACATACACATAC 150831
 OY 1802 ACTCAAGATCAGATGATCCTTATTTGGGCTTGGAAAAACAGGATG 1848
 DB 150830 ACTCAAGATCAGATGATCCTTATTTGGGCTTGGAAAAACAGGATG 150784
 RESULT 7
 AP003733 166867 bp DNA linear PRI 08-DEC-2001
 LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-810P12,
 DEFINITION complete sequence.
 ACCESSION AP003733
 VERSION AP003733.4 GI:17426128
 KEYWORDS HTG.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
 Homo sapiens genomic DNA
 TITLE Published Only in Database (2001)
 JOURNAL 2 (bases 1 to 166867)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
 TITLE Direct Submission
 JOINTL Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (SSC), Japan
 (E-mail:hattori@gs.riken.go.jp, URL:http://hgc.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 COMMENT On Dec 7, 2001 this sequence version replaced gi:17026124.
 FEATURES
 source
 1..166867
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-810P12"
 BASE COUNT 44705 a 38465 c 39268 g 44429 t
 ORIGIN
 Query Match 28.8%; Score 642.2; DB 9; Length 166867;
 Best Local Similarity 99.5%; Pred. No. 1e-139;
 Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1202 CAGCCTGAACAAGAGGATGGATTCAGCCCAATACAGAGAGAGAGATGCTCA 1261
 DB 12486 CAGCCTGAACAAGAGGATGGATTCAGCCCAATACAGAGAGAGAGATGCTCA 12545
 OY 1262 CGCTGGCATCATTTGGCGCTTCTAGGCTGCAGTCCCATGATCATTCTCCAGGCG 1321
 DB 12546 CGCTGGCATCATTTGGCGCTTCTAGGCTGCAGTCCCATGATCATTCTCCAGGCG 12605
 OY 1332 AAATCAAGACCAATCTAGTGTGGCCCAAGAGAGATCTCTTCCACAGAGGCGCTGCC 1381
 DB 12606 AAATCAAGACCAATCTAGTGTGGCCCAAGAGAGATCTCTTCCACAGAGGCGCTGCC 12665
 OY 1382 CAAAACCAAGAGAGGACCAAGACGTTAGGGGCGCCAGAAACCAAGGCTTGAA 1441
 DB 12666 CAAAACCAAGAGAGGACCAAGACGTTAGGGGCGCCAGAAACCAAGGCTTGAA 12735
 OY 1442 GCTTAAGGCTGTGAGCGCTTCAAGTCTGGCCCACTGATATCAGAGCCAGGCTACTACAG 1501
 DB 12726 GCTTAAGGCTGTGAGCGCTTCAAGTCTGGCCCACTGATATCAGAGCCAGGCTACTACAG 12785
 OY 1502 TGCCCAAGAGAGCGCCCTCAGGCCACATCCATGTTCTTCCCTAGAACATCAGCGCC 1561

DB 12786 TGCCCAAGAGAGCGCCCTCAGGCCACATCCATGTTCTTCCCTAGAACATCAGCGCC 12845
 OY 1562 GTCAAGGCTCAGAGTGTGACAGGATGACACCAAGACCAAGACCTTAAGACTGTGAG 1621
 DB 12846 GTCAAGGCTCAGAGTGTGACAGGATGACACCAAGACCAAGACCTTAAGACTGTGAG 12905
 OY 1622 TTCTGGGCGCAAGAAATTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTATGAGAGA 1681
 DB 12906 TTCTGGGCGCAAGAAATTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTATGAGAGA 12965
 OY 1682 CCCAGATATCTCAGATGAGAGAGAAACTGTGAGATTTAACTGACGAGATATGCCAGA 1741
 DB 12966 CCCAGATATCTCAGATGAGAGAGAAACTGTGAGATTTAACTGACGAGATATGCCAGA 13035
 OY 1742 GATCCCGAAATACCTCAAGAAACCTTTGGACATCACCACCAACATACACATAC 1801
 DB 13026 GATCCCGAAATACCTCAAGAAACCTTTGGACATCACCACCAACATACACATAC 13085
 OY 1802 ACTCAAGATCAGATGATCCTTATTTGGGCTTGGAAAAACAGGATG 1848
 DB 13086 ACTCAAGATCAGATGATCCTTATTTGGGCTTGGAAAAACAGGATG 13132
 RESULT 8
 AC004228 196080 bp DNA linear PRI 24-FEB-1999
 LOCUS Homo sapiens Chromosome 11q12.2 PAC clone pdu519013 containing
 DEFINITION human gene for ferritin heavy chain (FTH), complete sequence.
 ACCESSION AC004228
 VERSION AC004228.2 GI:4263838
 KEYWORDS HTG.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 196080)
 AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basil,M.,
 Buettner,J., Bumester,R., Card,P., desallibot,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
 HTGS Submission
 UNpublished
 TITLE 2 (bases 1 to 196080)
 JOURNAL Evans,G.A., Athanasiou,M., Basil,M., Bradbury,P., Brignac,S.,
 Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
 Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
 Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
 Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
 Ward,T. and Wilson,R.
 TITLE Direct Submission
 JOINTL Submitted (26-FEB-1998) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines
 Blvd, Dallas, TX 75235-8591, USA
 REFERENCE 3 (bases 1 to 196080)
 AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basil,M.,
 Buettner,J., Butler,C., Card,P., desallibot,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T.
 TITLE Direct Submission
 JOINTL Submitted (24-FEB-1999) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines
 Blvd, Dallas, TX 75235-8591, USA
 COMMENT
 On Feb 24, 1999 this sequence version replaced gi:2911733.
 IMPORTANT: This submission contains the entire insert of clone
 pdu519013. pdu519013 comes from the RPEC-3 PAC library constructed
 at the Roswell Park Cancer Institute by the Pletcher de Jong group.
 CHROMOSOMAL LOCUS: This PAC clone comes from the chromosome
 11p12.2 Best's disease region mapped between STS D11S461 and EST
 AHNMK. This region spans over 1.5 Mbp.

MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA), STSS
D11659 and WI-7524
MAPPED CLONE OVERLAP: PACs PDJ466a11 and PDJ756b9.

FEATURES
Location/Qualifiers
1. 196060

source
Repeat_region /organism="Homo sapiens"
Repeat_region /db_xref="taxon:9606"
Repeat_region /rpt_family="Alu" complement(2210..2482)
Repeat_region /rpt_family="Alu" 3457..3743
Repeat_region /rpt_family="Alu" complement(20145..20758)
Repeat_region /rpt_family="Alu" complement(23194..23599)
Repeat_region /rpt_family="Alu" complement(27367..27685)
Repeat_region /rpt_family="Alu" 28047..28457
Repeat_region /rpt_family="Alu" complement(29390..29565)
Repeat_region /rpt_family="Alu" complement(31076..31255)
Repeat_region /rpt_family="Alu" complement(34948..35045)
Repeat_region /rpt_family="Alu" 35232..35538
Repeat_region /rpt_family="Alu" complement(36242..36540)
Repeat_region /rpt_family="Alu" 37238..37522
Repeat_region /rpt_family="Alu" complement(38555..38650)
Repeat_region /rpt_family="Alu" complement(40483..41164)
Repeat_region /rpt_family="Alu" 42042..42116
Repeat_region /rpt_family="MIR" 42677..42817
Repeat_region /rpt_family="L1" 42856..43157
Repeat_region /rpt_family="Alu" 43148..43380
Repeat_region /rpt_family="L1" complement(43734..43999)
Repeat_region /rpt_family="L1" complement(46998..47169)
Repeat_region /rpt_family="MIR" complement(48048..48346)
Repeat_region /rpt_family="Alu" 48482..48737
Repeat_region /rpt_family="Alu" complement(48893..49037)
Repeat_region /rpt_family="L1" 49118..49292
Repeat_region /rpt_family="Alu" 49791..50059
Repeat_region /rpt_family="Alu" 50436..50501
Repeat_region /rpt_family="Alu" complement(51147..51441)
Repeat_region /rpt_family="Alu" complement(52700..52979)
Repeat_region /rpt_family="Alu" 53281..53568
Repeat_region /rpt_family="Alu" 54071..54204
Repeat_region /rpt_family="Alu" 54506..54793
Repeat_region /rpt_family="Alu" 54807..55061
Repeat_region /rpt_family="Alu" complement(64533..64826)
Repeat_region /rpt_family="Alu"

Repeat_region complement(65245..65783)
Repeat_region /rpt_family="Alu" complement(68465..68732)
Repeat_region /rpt_family="Alu" complement(70191..70476)
Repeat_region /rpt_family="Alu" 70560..70692
Repeat_region /rpt_family="Alu" 72294..72470
Repeat_region /rpt_family="Alu" complement(72800..73095)
Repeat_region /rpt_family="Alu" 73424..73695
Repeat_region /rpt_family="Alu" 73784..74071
Repeat_region /rpt_family="Alu" 74652..74935
Repeat_region /rpt_family="Alu" complement(75265..75514)
Repeat_region /rpt_family="Alu" complement(76086..76353)
Repeat_region /rpt_family="Alu" complement(78038..78312)
Repeat_region /rpt_family="Alu" 78513..78734
Repeat_region /rpt_family="Alu" 78866..79180
Repeat_region /rpt_family="Alu" complement(79384..79673)
Repeat_region /rpt_family="Alu" complement(81398..81690)
Repeat_region /rpt_family="Alu" 82013..82288
Repeat_region /rpt_family="Alu" complement(83419..83600)
Repeat_region /rpt_family="MIR" complement(87320..87624)
Repeat_region /rpt_family="Alu" complement(87710..87987)
Repeat_region /rpt_family="Alu" complement(88186..88472)
Repeat_region /rpt_family="Alu" 88822..89049
Repeat_region /rpt_family="Alu" 89224..89308
Repeat_region /rpt_family="MER20" 89729..90487
Repeat_region /rpt_family="Alu" 90504..90618
Repeat_region /rpt_family="Alu" complement(93313..93952)
Repeat_region /rpt_family="Alu" 96938..97517
Repeat_region /rpt_family="Alu" complement(104199..104514)
Repeat_region /rpt_family="Alu" complement(109738..109979)
Repeat_region /rpt_family="Alu" 110022..110141
Repeat_region /rpt_family="Alu" complement(111091..112016)
Repeat_region /rpt_family="Alu" 113210..113419
Repeat_region /rpt_family="L1" complement(114951..115005)
Repeat_region /rpt_family="MIR" 115685..115791
Repeat_region /rpt_family="MIR" 116870..117005
Repeat_region /rpt_family="MIR" complement(117877..118172)
Repeat_region /rpt_family="Alu" 118352..118600

repeat_region /rpl_family="Alu"
118679.118813
repeat_region /rpl_family="Alu"
complement(121233.121520)
repeat_region /rpl_family="Alu"
122658.123071

Query Match 28.8%; Score 642.2; DB 9; Length 196080;
Best Local Similarity 99.5%; Pred. No. 1.1e-139;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1202 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
DB 191325 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 191384
QY 1262 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCTCTCCAGAGGC 1321
DB 191385 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCTCTCCAGAGGC 191444
QY 1322 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGATCCCTTCCAGAGAGAGGCCTGCC 1381
DB 191445 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGAGGCCTGCC 191504
QY 1382 CAAAACCAAG 1441
DB 191505 CAAAACCAAG 191564
QY 1442 GCTTAAAGCTGTGAG 1501
DB 191565 GCTTAAAGCTGTGAG 191624
QY 1502 TGCCCAACAG 1561
DB 191625 TGCCCAACAG 191684
QY 1562 GTCAAAAGCTTACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
DB 191685 GTCAAAAGCTTACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191744
QY 1622 TTCTGGGGCCCAAG 1681
DB 191745 TTCTGGGGCCCAAG 191804
QY 1682 CCCAAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
DB 191805 CCCAAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191864
QY 1742 GATCCCCGAAATACACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1801
DB 191865 GATCCCCGAAATACACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191924
QY 1802 ACTCAAAAGATCAATGATGCTTATTTGGCCCTTGAGAGAGAGAGAGAGAGAGAGAG 1848
DB 191925 ACTCAAAAGATCAATGATGCTTATTTGGCCCTTGAGAGAGAGAGAGAGAGAGAGAG 191971

RESULT 9
HSYMD2P10 706 bp DNA linear PRI 22-OCT-1998
DEFINITION Homo sapiens vitelliform macular dystrophy protein (VMD2) gene,
exon 10.
ACCESSION AF073499
VERSION AF073499.1 GI:3598873
KEYWORDS
SEGMENT 10 of 11
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS Stohr,H., Marguardt,A., Rivera,A., Cooper,P.R., Nowak,N.J.,
Shows,T.B., Gerhard,D.S. and Weber,B.H.
A gene map of the Best's vitelliform macular dystrophy region in

JOURNAL chromosome 11q12-q13.1
MEDLINE Genome Res. 8 (1), 48-56 (1998)
PUBMED 98112782
REFERENCE 9445487
AUTHORS 2 (bases 1 to 706)
Marguardt,A., Stohr,H., Passmore,L.A., Kramer,F., Rivera,A. and
Weber,B.H.

TITLE Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(Best's disease)
Hum. Mol. Genet. 7 (9), 1517-1525 (1998)

JOURNAL mutations 1 to 706
MEDLINE 98367043
PUBMED 9700209

REFERENCE 3 (bases 1 to 706)
AUTHORS Marguardt,A., Stohr,H., Passmore,L., Kraemer,F., Rivera,A. and
Weber,B.H.F.
TITLE Direct Submission
Submitted (23-JUN-1998) Human Genetics, University, Biozentrum, Am
Hubland, Wurzburg 97074, Germany
FEATURES Location/Qualifiers

SOURCE

1. 706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/clone="PAC d466a11"
36..674
/gene="VMD2"
/number=10

BASE COUNT 204 a 209 c 164 g 129 t

ORIGIN

Query Match 28.7%; Score 640.6; DB 9; Length 706;
Best Local Similarity 99.4%; Pred. No. 1.1e-139;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1202 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
DB 33 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 92
QY 1262 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCTCTCCAGAGGC 1321
DB 93 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCTCTCCAGAGGC 152
QY 1322 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGATCCCTTCCAGAGAGAGGCCTGCC 1381
DB 133 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGAGGCCTGCC 212
QY 1382 CAAAACCAAG 1441
DB 213 CAAAACCAAG 272
QY 1442 GCTTAAAGCTGTGAG 1501
DB 273 GCTTAAAGCTGTGAG 332
QY 1502 TGCCCAACAG 1561
DB 333 TGCCCAACAG 392
QY 1562 GTCAAAAGCTTACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
DB 393 GTCAAAAGCTTACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
QY 1622 TTCTGGGGCCCAAG 1681
DB 453 TTCTGGGGCCCAAG 512
QY 1682 CCCAAGATATCTCAAGTGAAG 1741
DB 513 CCCAAGATATCTCAAGTGAAG 572
QY 1742 GATCCCCGAAATACACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1801

RESULT	10
AC003025	
LOCUS	
DEFINITION	Human Chromosome 11p12.2 PAC clone pDU166a11, complete sequence.
ACCESSION	AC003025
VERSION	AC003025.1 GI:3337308
KEYWORDS	HTG.
SOURCE	Homo sapiens.

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 112309)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buettner,J., Bumeister,R., Card,P., desallboast,F., Dunn,J.,
Englisch,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schegemen,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
HTGS Submission
Unpublished
2 (bases 1 to 112309)
Evans,G.A., Athanasiou,M., Bradbury,P., Brihanac,S., Bumeister,P.

TITLE	JOURNAL	REFERENCE	AUTHORS
Submitted (21-Oct-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA		3 (bases 1 to 112309)	Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Pasit, M.

English, C., Ehrhridge, S., Garner, H. R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K. N., Major, T., McFarland, J., Newton, C., Osborne-Lawrence, S., Schageman, J., Schultz, R. A., Stimson, S., Syed, M. and Ward, P.
Direct Submission
Submitted (23-JUL-1998) Genome Science & Technology Center
JOURNAL

COMMENT
Blvd, Dallas, TX 75235-8591, USA
On Jul 23, 1998 this sequence version replaced gi:2554967.
IMPORTANT: This submission contains the entire insert of clone
PDD2466a11. PDD2466a11 comes from a PAC library constructed at the
Roswell Park Cancer Institute by the Pletcher de Jong group. This
clone has been finished according to strict quality criteria and
attempts have been made to resolve all base calling problems such
as compressions and repetitive elements. The expected Phred/Phrap
calculated errors/10kb is 0.18. In addition, this sequence has
been finished such that 99.9% of consensus base calls consist of
either double stranded coverage or 2 types of labeling chemistry on
one strand.

Annotation of pjd466a1 can be found at <http://gestec.swmed.edu/chromosol.htm>. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between SRS D11s61 and EST AHNAK. This region spans over 1.5 Mbp. MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA). MAPPED CLONE OVERLAP: HTGS, submitted PAC clones pjd159c3 and pjd756b9.

source 1. .112309
/organism="Homo sapiens"

repeat_region	/db_xref=taxon:9606"	
repeat_region	/rpt_family="Alu"	.4656)
repeat_region	/rpt_family="Alu"	.10130)
repeat_region	/rpt_family="Alu"	.10163 .10282
repeat_region	/rpt_family="Alu"	complement(11232 .
repeat_region	/rpt_family="Alu"	13351 .13560
repeat_region	/rpt_family="L1"	complement(15092 .
repeat_region	/rpt_family="MIR"	15826 .15932
repeat_region	/rpt_family="MIR"	17011 .17146
repeat_region	/rpt_family="MIR"	complement(18018 .
repeat_region	/rpt_family="Alu"	18493 .18741
repeat_region	/rpt_family="Alu"	18820 .18954
repeat_region	/rpt_family="Alu"	complement(21374 .
repeat_region	/rpt_family="Alu"	22805 .23212
repeat_region	/rpt_family="Alu"	23499 .23752
repeat_region	/rpt_family="Alu"	24118 .24472
repeat_region	/rpt_family="Alu"	25005 .25113
repeat_region	/rpt_family="MIR"	31160 .31442
repeat_region	/rpt_family="Alu"	complement(32651 .
repeat_region	/rpt_family="MIR"	complement(33570 .
repeat_region	/rpt_family="MIR"	38958 .39145
repeat_region	/rpt_family="Alu"	44878 .45480
repeat_region	/rpt_family="Alu"	49763 .50071
repeat_region	/rpt_family="Alu"	complement(50172 .
repeat_region	/rpt_family="Alu"	50902 .50982
repeat_region	/rpt_family="MIR"	51038 .51659
repeat_region	/rpt_family="Alu"	54494 .54769
repeat_region	/rpt_family="Alu"	55529 .55623
repeat_region	/rpt_family="THE1"	complement(55691 .
repeat_region	/rpt_family="Alu"	complement(56049 .
repeat_region	/rpt_family="Alu"	56721 .57100
repeat_region	/rpt_family="THE1"	58402 .58680
repeat_region	/rpt_family="Alu"	complement(58748 .
repeat_region	/rpt_family="Alu"	59678 .59904
repeat_region	/rpt_family="Alu"	61149 .61424
repeat_region	/rpt_family="Alu"	complement(61930 .
repeat_region	/rpt_family="Alu"	62276 .62710
repeat_region	/rpt_family="MUT1"	

```
repeat_region 62915..63118
                /rpt_family="Alu"
repeat_region complement(63349)..63455)
                /rpt_family="Alu"
repeat_region 64339..64614
                /rpt_family="Alu"
repeat_region complement(64707)..64984)
                /rpt_family="Alu"
repeat_region complement(65596)..65811)
                /rpt_family="Alu"
repeat_region complement(66500)..66753)
                /rpt_family="Alu"
repeat_region 67311..67542
                /rpt_family="Alu"
repeat_region 67835..68119
                /rpt_family="Alu"
repeat_region 68860..69150
                /rpt_family="Alu"
repeat_region 69164..69272
                /rpt_family="Alu"
repeat_region 69354..70245
                /rpt_family="Alu"
repeat_region 70655..70939
                /rpt_family="Alu"
repeat_region 70933..71075
                /rpt_family="L1"
repeat_region 71119..71395
                /rpt_family="Alu"
repeat_region 71404..71530
                /rpt_family="LTR5"
repeat_region 71536..71637
                /rpt_family="LTR5"
repeat_region 72720..72948
                /rpt_family="L1"
repeat_region complement(73488)..73633)
                /rpt_family="Alu"
repeat_region complement(73925)..74200)
                /rpt_family="Alu"
repeat_region 75721..75955
                /rpt_family="Alu"
repeat_region 77185..77449
                /rpt_family="Alu"
repeat_region 78093..78222
                /rpt_family="L1"
repeat_region 78327..78294
                /rpt_family="L1"
repeat_region complement(80458)..80707)
                /rpt_family="Alu"
repeat_region complement(81544)..81974)
                /rpt_family="Alu"
repeat_region 82055..82305
                /rpt_family="Alu"
repeat_region 82514..83096
                /rpt_family="Alu"
repeat_region complement(83685)..83940)
                /rpt_family="Alu"
repeat_region 85336..85850
                /rpt_family="Alu"
repeat_region complement(86429)..86483)
                /rpt_family="MIR"
repeat_region 86902..87183
                /rpt_family="Alu"
repeat_region 87546..87792
                /rpt_family="Alu"
```

Query Match 28.7% Score 640.6; DB 9; Length 112309;
Best Local Similarity 99.4%; Pred. No. 2.3e-139;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1202 CAGCCTGACAAAGAGAGATGGAGTTCAGCCCAATCAGAGAGACGAGAGATGCTCA 1261
DB 91470 CAGCCTGACAAAGAGAGATGGAGTTCAGCCCAATCAGAGAGACGAGAGATGCTCA 91529

```
QY 1262 CGCTGCGATCATTTGGCGCTTCTCTAGGCGCTGAGTCCATGATCATCCTCCAGGGC 1321
                |||
Db 91530 CGCTGCGATCATTTGGCGCTTCTCTAGGCGCTGAGTCCATGATCATCCTCCAGGGC 91589
QY 1322 AAATCTCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACGAGGGCTCC 1381
                |||
Db 91590 AAATCTCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACGAGGGCTCC 91649
QY 1382 CAAAACCCACAGGCGACCAAAAGAACGTTAGGCGCCAGAGACAAACAGGCTCGAA 1441
                |||
Db 91650 CAAAACCCACAGGCGACCAAAAGAACGTTAGGCGCCAGAGACAAACAGGCTCGAA 91709
QY 1442 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGGCCACTGTATGAGAGGCGGCTACTAG 1501
                |||
Db 91710 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGGCCACTGTATGAGAGGCGGCTACTAG 91769
QY 1502 TGCCCAACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGGCC 1561
                |||
Db 91770 TGCCCAACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGGCC 91829
QY 1562 GTCAAGCTTACAGTGTACAGGCAATGACACCAACAAAGCTTAAAGACTGTAG 1621
                |||
Db 91830 GTCAAGCTTACAGTGTACAGGCAATGACACCAACAAAGCTTAAAGACTGTAG 91889
QY 1622 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTACAGAGCGATGGGCGCTGATGAGCA 1681
                |||
Db 91890 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTACAGAGCGATGGGCGCTGATGAGCA 91949
QY 1682 CCCAAGATATCTCAAGTACAGAGAGAAACGTGTGAGTTTAACTGACGATATGCGAGA 1741
                |||
Db 91950 CCCAAGATATCTCAAGTACAGAGAGAAACGTGTGAGTTTAACTGACGATATGCGAGA 92009
QY 1742 GATCCCGGAAATATACCTCCTCAAGAAACCTTGGAAACAACACCAACATACACATAC 1801
                |||
Db 92010 GATCCCGGAAATATACCTCCTCAAGAAACCTTGGAAACAACACCAACATACACATAC 92069
QY 1802 ACTCAAGATCATCATGATCCTTATTTGGGCGCTTGGAAACAGAGGATG 1848
                |||
Db 92070 ACTCAAGATCATCATGATCCTTATTTGGGCGCTTGGAAACAGAGGATG 92116
```

RESULT 11
AC084857 133683 bp DNA linear HTG 22-NOV-2000
LOCUS Homo sapiens chromosome 11, clone CTD-3231N5 map 11, WORKING DRAFT
DEFINITION
ACCESSION AC084857
VERSION AC084857.1 GI:11276215
KEYWORDS HTG; HTGS; PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 133683)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campilano,A., Castle,A.,
Choapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeRubeis,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heath,J., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karats,A., LaRoque,K.,
Lamasares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McNeeters,R., Meldrum,J., Menus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,

TITLE
JOURNAL
COMMENT

Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J., Zimmer, A., and Zody, M.

Direct Submission

Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L11038

Center clone name: 3231_N-5

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator; Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 126761 bases at least Q40

Consensus quality: 129952 bases at least Q30

Consensus quality: 131351 bases at least Q20

Insert size: 128000; agarose-fp

Insert size: 132483; sum-of-contigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 5.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1459: contig of 1459 bp in length
1460 1559: gap of 100 bp
1560 2520: contig of 961 bp in length
2521 2620: gap of 100 bp
2621 6334: contig of 3714 bp in length
6335 6434: gap of 100 bp
6435 11913: contig of 5479 bp in length
11914 12013: gap of 100 bp
12014 20756: contig of 8743 bp in length
20757 20856: gap of 100 bp
20857 29735: contig of 8879 bp in length
29736 29835: gap of 100 bp
29836 38645: contig of 8810 bp in length
38646 38745: gap of 100 bp
38746 52193: contig of 13448 bp in length
52194 52293: gap of 100 bp
52294 65209: contig of 12916 bp in length
65210 65309: gap of 100 bp
65310 79796: contig of 14487 bp in length
79797 79896: gap of 100 bp
79897 95896: contig of 16000 bp in length
95897 95996: gap of 100 bp
95997 113419: contig of 17423 bp in length
113420 113519: gap of 100 bp
113520 133683: contig of 20164 bp in length.

```

FEATURES

SOURCE

```

1. 133683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-3231N5"
/clone_lib="CTD Human BAC"
misc_feature
1..1459

```

```

misc_feature /note="assembly-fragment"
1560..2520
/note="assembly-fragment"
2621..6334
misc_feature /note="assembly-fragment"
6435..11913
/note="assembly-fragment"
12014..20756
misc_feature /note="assembly-fragment"
20857..29735
/note="assembly-fragment"
29836..38645
misc_feature /note="assembly-fragment"
38746..52193
/note="assembly-fragment"
52294..65209
misc_feature /note="assembly-fragment"
65310..79796
/note="assembly-fragment"
79897..95896
misc_feature /note="assembly-fragment"
95997..113419
/note="assembly-fragment"
113520..133683
/note="assembly-fragment"

BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
ORIGIN
Query Match 28.7% Score 640.6; DB 2; Length 133683;
Best Local Similarity 99.4%; Pred. No. 2.4e-139;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1202 CAGCCTGACAAAGAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
Db 92312 CAGCCTGACAAAGAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 92371

QY 1262 CGCTGGCATATTTGGCCGCTTCTAGGCTGACGCTCCATGATACCATCTCCAGGCG 1321
Db 92372 CGCTGGCATATTTGGCCGCTTCTAGGCTGACGCTCCATGATACCATCTCCAGGCG 92431

QY 1322 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCCACGAGGCGCTGCC 1381
Db 92432 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCCACGAGGCGCTGCC 92491

QY 1382 CAAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441
Db 92492 CAAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 92551

QY 1442 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGGCTACTACAG 1501
Db 92552 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGGCTACTACAG 92611

QY 1502 TGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
Db 92612 TGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 92671

QY 1562 GTCAAAGCTTCACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
Db 92672 GTCAAAGCTTCACAGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 92731

QY 1622 TTCTGGGGGCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
Db 92732 TTCTGGGGGCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 92791

QY 1682 CCCAGAGATCTCAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
Db 92792 CCCAGAGATCTCAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 92851

QY 1742 GATCCCGGAAATCACTCTCAAGAACTTTGGAACTATCAGACCAACCAATACACACTAC 1801
Db 92852 GATCCCGGAAATCACTCTCAAGAACTTTGGAACTATCAGACCAACCAATACACACTAC 92911

QY 1802 ACTCAAGATCATGATGATCTTATTTGGGCTTGGAAAGAGAGATG 1848

```


Db 92912 ACTCAAGATCATGATCCTTATGTGGCTTGGAAGAGGTCTG 92958

RESULT 12
AF440756 1908 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens vitelliform macular dystrophy 2-like protein 1 mRNA,
DEFINITION complete cds.
ACCESSION AF440756
VERSION AF440756.1 GI:21734839
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1908)
AUTHORS Marguardt,A., Stohr,H., Passmore,L.A., Kramer,F., Rivera,A. and
Weber,B.H.
TITLES Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(Best's disease)
JOURNAL Hum. Mol. Genet. 7 (9), 1517-1525 (1998)
MEDLINE 98367043
PUBMED 9702029
REFERENCE 2 (bases 1 to 1908)
AUTHORS Stohr,H., Marguardt,A., Nanda,I., Schmid,M. and Weber,B.H.
TITLES Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family
JOURNAL Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 12032749
PUBMED 12032738
REFERENCE 3 (bases 1 to 1908)
AUTHORS Stohr,H., Marguardt,A. and Weber,B.H.F.
TITLES Three novel human VMD2-like protein genes are members of the
evolutionary highly conserved RFP family
JOURNAL Unpublished
AUTHORS 4 (bases 1 to 1908)
TITLES Stohr,H., Marguardt,A. and Weber,B.H.F.
REFERENCE Direct Submission
AUTHORS Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg,
JOURNAL Biozentrum, Am Hubland, Wuerzburg 97074, Germany
FEATURES
source location/Qualifiers
1..1908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.12-p13.2"
1..1530
/note="VMD2L1: putative membrane protein; Pfam01062;
worm_family_8"
/codon_start=1
/product="vitelliform macular dystrophy 2-like protein 1"
/protein_id="AA076995.1"
/db_xref="GI:21734840"
/translation="MTVTYTAAYANARFGFSQLLLMRGSITKLLMRELLCFLGYM
ALSAAYRFLVTEGQKRYFERLYVCDQYSLIPVSFLGTYVLVVRMMSOYLCHPL
PALMKVAGTVGRDGRGLRYRLTLARYAGLSAVILRSVAVRERPTIDHYEA
GEMTREERKKEFNLSYNKYVPCWFSLSLAOARREGIRNSALKILLELVNFR
GKGMLEHFDWISVPLVYQVYTAIXSYLAFLIGROPIDPAGOKKDDHLDLCVIF
TILQFFRYGWLKVAEQLINPREGDDDEFTNFIIDNRFVSMIAVEMDDIAVLEK
DLWDAAEARAPYTAATVAFQLRQPSFGSTFDTTLAKEDMQFRLDGLDPMGEABGD
FLORLIPAGAGVAGGPIGLRSLFLRKNSCVSEASTGASCAYVPEGAAPEACSGD
PLIDPLPEPEAPAPGAPBPLTLIPGVEFEFSIVTMPGPRGAPMLPSIGEEENL
A"

BASE COUNT 337 a 604 c 551 g 416 t

Query Match 25.9%; Score 576.8; DB 9; Length 1908;
Best Local Similarity 70.0%; Pred. No. 1.3e-124;
Matches 792; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

QY 105 ATGACATCACTTACACCAAGCGCTAATGCCCGCTTACGCTCTTCCCGCTG 164

Db 1 ATGACGCTACCTTACACAGCCGAGTGGCCAGACGCCGCTTCGGTGTCTCCAGCTG 60

QY 165 CTGCTGTGTCGCGGGCAGCATCTACAGCTGCTATATGGAGTTCTTATCTCTG 224

Db 61 CTGCTACTGTGCGCTGGAGCATCTACAACTCTGTGGAGAGCTCTGCTCTT 120

QY 225 CTGCTACTGTGCGCTGGAGCATCTACAACTCTGTGGAGAGCTCTGCTCTT 284

Db 121 GGGTCTACATGCGCGTGAAGTGTGCTTACCGCTTGTGTGACCGAAGGCGCAAGCGC 180

QY 285 ATGTTGAGAAACAGACTGTGTATGGCAGCTACATCCAGCTCATCCCATTTCTTC 344

Db 181 TACTTCAGAAAGCTGTGTATTTATGTGACCAAGTATGCAACCTCATCTCTCTTC 240

QY 345 GTGCTGGCTTCTACGTGACGCTGTGTGACCCGCTGTGTGGAACAGTACAGAACTG 404

Db 241 GTGCTTGGCTTCTTATGTGACGCTGTGTGGAACCGTGTGTGAGACCACTTATGATG 300

QY 405 CCGTGGCCGCGCCGCTCATGAGCTGTGTGCGGGCTTGTGGAAGCAAGAGAGCA 464

Db 301 CCGTGGCCGCGCGCTCATGAGCTGTGTGCGGGCTTGTGGAAGCAAGAGAGCA 360

QY 465 GCGCGCTGTGCGCGCCAGCTCATCGCTACGCCCAACCTGGCCAACTGCTCATCTG 524

Db 361 GCGCGCTGTACGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 420

QY 525 GCGAGCTGTACGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 584

Db 421 GCGCGCTGTACGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 480

QY 585 GCGTGTATGATCGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 644

Db 481 GCGTGTATGATCGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 540

QY 645 TTCTGGTGTGCGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 704

Db 541 TACTGGTGTGCGCGCGCCAGCTCATCGCTACGCCCAAGGCTGTGCGCGCTCATCTG 600

QY 705 ATCCGGGACCTTATCCGCTCCAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 764

Db 601 ATCCGGGACCTTATCCGCTCCAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 660

QY 765 TGTGACACCTGTATGCTACGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 824

Db 661 TGTGACACCTGTATGCTACGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 720

QY 825 ACTGTGGCTGTACGCTTCTCTGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 884

Db 721 ACCATCGCAGCTGTACGCTTCTCTGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 780

QY 885 GCCAAGGCTACCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAG 944

Db 781 GCTCAGGCTTACAAAGCAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAG 840

QY 945 TTCTTCTTATGTTGCTGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 1004

Db 841 TTCTTCTTATGTTGCTGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 900

QY 1005 GATGATGATATTTTGTGAGCAACTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 1064

Db 901 GAGATGATGATATTTTGTGAGCAACTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 960

QY 1065 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124

Db 961 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

QY 1125 CCGGAGCCAGACCCGCTTACAGAGCTGTGCTTCCGCTGTGAGAGCTGTGAGAGCT 1181

Db 1021 GCGGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTG 1080

QY 1182 ATGGGCTCAGCTTCAACATCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1233


```

misc_feature      86594..96688
                  /note="assembly_name:Contig36"
misc_feature      96789..107169
                  /note="assembly_name:Contig37"
misc_feature      107270..117047
                  /note="assembly_name:Contig38"
misc_feature      117148..127120
                  /note="assembly_name:Contig39
                  clone_end:SP6
                  vector_side:right"
misc_feature      127221..142029
                  /note="assembly_name:Contig40"
misc_feature      142130..160169
                  /note="assembly_name:Contig41"
BASE COUNT      41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN

```

```

Query Match      24.8%; Score 552.8; DB 2; Length 160169;
Best Local Similarity 99.3%; Pred. No. 1.1e-118;
Matches 576; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1202 CACCTGACAAAGAGAGAGATGAGATTCAGCCCATCAGAGAGAGAGAGATGCTCA 1261
      |||||
DB 67594 CACCTGAAACAAAGAGAGATGAGATTCAGCCCATCAGAGAGAGAGAGATGCTCA 1261
      |||||

QY 1262 CGCTGGCATCTTGGCCGCTTCTAGGCTGAGATGCCATGATCAGATCCCTCCAGGGC 1321
      |||||
DB 67535 CGCTGGCATCTTGGCCGCTTCTAGGCTGAGATGCCATGATCAGATCCCTCCAGGGC 1321
      |||||

QY 1322 AAACCTCAAGGACCAAACTACTGTGGCCCAAGAGAGGATCCCTTCTCCAGAGGCTGCC 1381
      |||||
DB 67475 AAACCTCAAGGACCAAACTACTGTGGCCCAAGAGAGGATCCCTTCTCCAGAGGCTGCC 1381
      |||||

QY 1382 CAAAAACCAAGGAGGAGCCAAAGAACTTAAAGGGCCAGAGAGAGAGAGAGAGAGAGAG 1441
      |||||
DB 67416 CAAAAACCAAGGAGGAGCCAAAGAACTTAAAGGGCCAGAGAGAGAGAGAGAGAGAGAG 1441
      |||||

QY 1442 GCTTAAGGCTGTGGAGGCTTCAAGTGTGGCCCAAGAGAGGAGGCTACTACAG 1501
      |||||
DB 67356 GCTTAAGGCTGTGGAGGCTTCAAGTGTGGCCCAAGAGAGGAGGCTACTACAG 1501
      |||||

QY 1502 TGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1561
      |||||
DB 67296 TGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1561
      |||||

QY 1562 GTCAAGGCTTCAAGTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1621
      |||||
DB 67236 GTCAAGGCTTCAAGTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1621
      |||||

QY 1622 TTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1681
      |||||
DB 67176 TTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1681
      |||||

QY 1682 CCCAGAGATCTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1741
      |||||
DB 67116 CCCAGAGATCTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1741
      |||||

QY 1742 GATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
      |||||
DB 67056 GATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
      |||||

```

```

RESULT 14
BC031186      1956 bp      mRNA      linear      ROD 07-AUG-2002
LOCUS      BC031186      Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
DEFINITION      BC031186
ACCESSION      BC031186.1 GI:21411099
VERSION      MGC.
KEYWORDS      house mouse.
SOURCE      Mus musculus
ORGANISM      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE      1 (bases 1 to 1956)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mdc@paxill.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRMA Plate: 58 Row: 9 Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

```

source      1..1956
            /location/Qualifiers
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /map="FVB/N"
            /clone="MGC:37621 IMAGE:4989959"
            /tissue_type="Colon, normal, 5 month old male mouse."
            /clone_id="NCI CGAP_C024"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"
            /note="1573
            /codon_start=1
            /product="unknown (protein for MGC:37621)"
            /protein_id="AAH31186.1"
            /db_xref="GI:21411100"
            /db_xref="LocusID:212989"
            /translation="MALSAARFLAEERKRYEFLVYCDQVSLIPSVFLGYYVT
            LVHRRMNOYLCPMLPDLMLCIVAGTVHGRDRRLRYRTLMRAGLSAVILRSVST
            AVFKRPFLIDHVRAGFTREKREKFLNSYKMYPCVWFSLSLAQIGROFLDPA
            NSALKLLEHNVNRSKCMGFHVDWISIPYVOYVIAVSYFLAQIGROFLDPA
            OGKRDHFLDCLVPIFTLLQFFYAGMLKVAQQLINPREDDDDEPTNLLDRNQVSM
            LAVDEMIDDLAMEKEDLYMDAEARAPFYTAFLDLOPSFGSTFDIALAKEDMOFO
            RLDDVDPLGEVHGDFLQRLPLPAGAGSVGPRSLILRRNSVSEASTVASCAG
            AADGGVCEGCGDLDPLPSLEPELEPACPEPAPJPGPTPEPTTVSIPGPAPAP
            PWLSPFGEESPSA"
            /cds

```

```

BASE COUNT      411 a 577 c 530 g 438 t
ORIGIN
Query Match      24.2%; Score 538.4; DB 10; Length 1956;
Best Local Similarity 67.6%; Pred. No. 1.4e-115;
Matches 771; Conservative 0; Mismatches 366; Indels 3; Gaps 1;

```

```

QY 97 GCTTGCCATGACCATCTTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 156
      |||||
DB 39 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 156
      |||||

QY 157 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 216
      |||||
DB 99 GCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
      |||||

QY 217 TCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276
      |||||
DB 159 GTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276
      |||||

QY 277 AACAGCTGATTTGAGAGAACTGACTGCTGATTTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 336
      |||||
DB 219 AGAAGCGCTACTGAGAGAGCTGTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336
      |||||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:51:03 ; Search time 6617.31 Seconds
(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaaattgtctctctg.....tttgagagcgagcgaggt 5000

Scoring table: IDENTITY NUC
Gapop: 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hlc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hlc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_man:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	546	10.9	2453	11	BC015220 Homo sapi
2	541.8	10.8	545	17	A0588060 CITR1-EI-
3	434	8.7	448	9	A1125626 q994f08.x
4	433	8.7	451	9	A1808844 wF97f09.x
5	431	8.6	437	9	A1432176 tg77f09.x
6	428.8	8.6	921	12	BE785274 601474631

7	424	8.5	442	9	A1028274	A1028274 ov84e04.x
8	387.4	7.7	423	9	A1432177	A1432177 tg77f10.x
9	296.8	5.9	2971	17	AF101960	AF101960 AF101960
10	293.8	5.9	2971	17	AF101960	AF101960 AF101960
11	272.2	5.4	1641	11	BC033224	BC033224 Homo sapi
12	263.8	5.3	590	17	A0588588	A0588588 CITR1-EI-
13	263.8	5.3	814	14	B0440115	B0440115 AGENCOURT
14	261	5.2	261	9	AA921904	AA921904 cm04f05.8
15	257.6	5.2	559	9	AU148047	AU148047 AU148047
16	257.4	5.1	652	9	AU120416	AU120416 AU120416
17	257.2	5.1	821	12	BE974661	BE974661 BE974661
18	255	5.1	658	17	AG079111	AG079111 Pan trogl
19	254.2	5.1	515	14	B0581726	B0581726 112c10.x
20	254.2	5.1	539	17	A0543001	A0543001 RPCI-11-3
21	253.4	5.1	698	14	B0949414	B0949414 ACENCOURT
22	253.4	5.1	698	14	B0709774	B0709774 UI-H-DIO-
23	252.8	5.1	661	17	AG114592	AG114592 Pan trogl
24	252.6	5.0	661	17	AF188083	AF188083 UP-479-24
25	252.4	5.0	705	17	AG059506	AG059506 Pan trogl
26	250	5.0	705	17	AQ346434	AQ346434 RPCI11-10
27	249.6	5.0	652	9	AU120416	AU120416 AU120416
28	248.4	5.0	481	9	AT791227	AT791227 oe21g10.y
29	248.4	5.0	575	9	AL601708	AL601708 DRFZp13A
30	247.8	5.0	643	17	A2515923	A2515923 RPCI-11-4
31	247.2	4.9	942	14	B0707956	B0707956 AGENCOURT
32	246.2	4.9	942	14	B0181755	B0181755 UI-H-EDU-
33	246	4.9	676	14	B0181755	B0181755 UI-H-EDU-
34	245.4	4.9	654	9	AU117926	AU117926 AU117926
35	245.4	4.9	1358	13	BM564410	BM564410 AGENCOURT
36	245	4.9	889	13	B1457655	B1457655 603198377
37	244	4.9	867	17	A0782080	A0782080 HS-3174.B
38	243.8	4.9	687	17	A0741949	A0741949 HS-5569.B
39	243.4	4.9	529	12	BG402924	BG402924 60218727
40	243.4	4.9	655	9	AL580585	AL580585 AL580585
41	243.4	4.9	778	12	BR663240	BR663240 602144358
42	243.4	4.9	947	12	BG680848	BG680848 602628648
43	243.4	4.9	1085	13	BM476343	BM476343 AGENCOURT
44	243.2	4.9	877	17	A0739838	A0739838 HS-5505.A
45	243	4.9	494	17	AQ180762	AQ180762 HS-3220.A

ALIGNMENTS

RESULT 1	BC015220	2453 bp	mRNA	Linear	HTC 25-JUL-2002
LOCUS	BC015220				
DEFINITION	Homo sapiens, clone IMAGE:3877806, mRNA.				
ACCESSION	BC015220				
VERSION	BC015220.1	GI:21955361			
KEYWORDS	HTC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 2453)				
TITLE	Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
JOURNAL	Direct Submission				
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				
COMMENT	Contact: MGC help desk				
	Email: cgabs-remail.nih.gov				
	Tissue Procurement: DCTD/DRP/Gazdar				
	CDNA Library Preparation: Life Technologies, Inc.				
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)				
	Center, Stanford University School of Medicine, Stanford, CA 94305				
	Web site: http://www-sngc.stanford.edu				
	Contact: (Dickson, Mark) mcd@pax11.stanford.edu				
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,				

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>
Series: IRAC Plate: 14 Row: b Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: Incomplete processing.

FEATURES

SOURCE

Location/Qualifiers

1. .2453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3877806"
/issue_type="lung, large cell carcinoma"
/clone_lib="NIH_MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT 633 a 711 c 606 g 503 t

ORIGIN

Query Match

Best Local Similarity 10.9%; Score 546; DB 11; Length 2453;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1037 GCAAGGCTGGGACCAAGAACAGGAGCTGTGACGTGACGAGCCGCGTATTCATTCTTCCAP 1096

12 GCAGGCTGGGACCAAGAACAGGAGCTGTGACGTGACGAGCCGCGTATTCATTCTTCCAP 71

1097 GCCACAGAGGCTGTCAAGAACGAGGCTAGTCAGAGGCTCTCTCTCTGAGAGT 1156

72 GCCACAGAGGCTGTCAAGAACGAGGCTAGTCAGAGGCTCTCTCTCTGAGAGT 131

1157 CCTGCGACAGAAAGTTGAAGCTCAGCAGAGCCCTTACCCTTCTCTCTGCAAGGC 1216

132 CCTGCGACAGAAAGTTGAAGCTCAGCAGAGCCCTTACCCTTCTCTCTGCAAGGC 191

1217 TCAGGGGTCAGAACACTGTGAGAGAGATCTTTAGCTTGGATTTAGGGCCATGGTA 1276

192 TCAGGGGTCAGAACACTGTGAGAGAGATCTTTAGCTTGGATTTAGGGCCATGGTA 251

1277 GAGGGGGTGTGCTTAAATTCAGACCCCTGCTCTCAGCCCAACCTTCCAGAGAAAT 1336

252 GAGGGGGTGTGCTTAAATTCAGACCCCTGCTCTCAGCCCAACCTTCCAGAGAAAT 311

1337 TAGAGGGGCTATGGCAGAGCTGTGCTAGCCGTTCTTGTGACAGATTACAAAGAGGAC 1396

312 TAGAGGGGCTATGGCAGAGCTGTGCTAGCCGTTCTTGTGACAGATTACAAAGAGGAC 371

1397 TTAAGCAAGAGACTCTTTGTGAGAGTCTGGGCTTGGAGTGAAGGAGGAGGCTCAGC 1456

372 TTAAGCAAGAGACTCTTTGTGAGAGTCTGGGCTTGGAGTGAAGGAGGAGGCTCAGC 431

1457 ACTCAGTGGGAGTCCAGGCTCTTAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 1516

432 ACTCAGTGGGAGTCCAGGCTCTTAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 491

1517 GGAATCCACCAAGCTTATGCGCAGAGCTTGTGTGGATCATGGAACCACTGGAAGCC 1576

492 GGAATCCACCAAGCTTATGCGCAGAGCTTGTGTGGATCATGGAACCACTGGAAGCC 551

1577 CACCTG 1582

552 CACCTG 557

RESULT 2

A0588060

LOCUS

DEFINITION

ACCESSION

VERSION

A0588060 545 bp DNA linear GSS 07-JUN-1999
CITR1-El Homo sapiens genomic clone 2643K4, DNA
sequence.
A0588060
A0588060.1 GI:5014740

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

LOCATION/Qualifiers

GSS.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Other GSS: CITR1-El-2643K4.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

SOURCE

LOCATION/Qualifiers

1. .545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2643K4"
/clone_lib="CITR1-El"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11, Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT

133 a 151 c 150 g 111 t

ORIGIN

Query Match

Best Local Similarity 9.6%; Score 541.8; DB 17; Length 545;

Matches 543; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

927 TGTCTATTACTAGGCTGATGAATTCACAGCAACACCATCTTTTCAGATTAAGGGCAC 986

1 TGTCTATTACTAGGCTGATGAATTCACAGCAACACCATCTTTTCAGATTAAGGGCAC 60

987 TAGGCTGAGAGAGAGCTGAATCTACCCGGGGGTACACACACAGGTGGCAAGGCTGG 1046

61 TAGGCTGAGAGAGAGCTGAATCTACCCGGGGGTACACACACAGGTGGCAAGGCTGG 120

1047 GACCAAGAACAGGAGCTGTGACTGACGAGCCGGTATTCATTTCTTCATAGCCACAGGG 1106

121 GACCAAGAACAGGAGCTGTGACTGACGAGCCGGTATTCATTTCTTCATAGCCACAGGG 180

1107 CTGTCAAGAGCCCAAGGGGCTAGTGAAGGCTCTCTCTCTGAGAGATTCTGACACAG 1166

181 CTGTCAAGAGCCCAAGGGGCTAGTGAAGGCTCTCTCTCTGAGAGATTCTGACACAG 240

1167 AAGTTGAAGCTAGCAGACAGCCCTTAACCCCAACTCTCTCTGCAAGGCTCAGGGGCTCA 1226

241 AAGTTGAAGCTAGCAGACAGCCCTTAACCCCAACTCTCTCTGCAAGGCTCAGGGGCTCA 300

1227 GAAACAGTGGAGAGAGATCTTTAGCTTGTGATTTTAAAGGCGCATGTAGAGGGGTGT 1286

301 GAAACAGTGGAGAGAGATCTTTAGCTTGTGATTTTAAAGGCGCATGTAGAGGGGTGT 360

1287 TGCCCTTAATTCAGAGCCCTGTCTCAGCCCAACACCTCCAAAGAAATTAAGAGGGCC 1346

361 TGCCCTTAATTCAGAGCCCTGTCTCAGCCCAACACCTCCAAAGAAATTAAGAGGGCC 420

1347 ATGGCCAGGCTGTGCTAGCCGTTGCTTCTGAGAGAGATTACAAAGAGGACTAAGACAAG 1406

421 ATGGCCAGGCTGTGCTAGCCGTTGCTTCTGAGAGAGATTACAAAGAGGACTAAGACAAG 480

OY		1407	ACTCCTTGTGAGAGTCCTCGCTTAGGAGATCAAGTGACGGCGCTGACGACTGCACGNG	1468
DB		481	ACTCCTTTGTGAGAGTCCTCGCTTAGGAGATCAAGTGACGGCGCTGACGACTGCACG	540
OY		1467	GCACT 1471	
DB		541	GCAGT 545	
RESULT 3				
A1125626				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
Source				
Query Match				
Best Local Similarity				
Matches 437; Conservative				
OY		2176	TGCTTCACTAATTATTTTATATGAGCCCTTCTACGAAACACAGAAGAGAGCTTCATCTG	2235
DB		7	TGCTTCACTAATTATTTTATATGAGCCCTTCTACGAAACACAGAAGAGAGCTTCATCTG	66
OY		2236	AGGAGGAAGACAGGCGAAGACAGGAGATCTGTATTAATTTCAAGTAGTGAATAAGTC	2295
DB		67	AGGAGGAAGACAGGCGAAGACAGGAGATCTGTATTAATTTCAAGTAGTGAATAAGTC	126
OY		2296	TCTCTAGAAAATATCAAGCAAGGTGAGAGACACAGACGCCGGTGGCAGTGGGCTCAT	2355
DB		127	TCTCTAGAAAATATCAAGCAAGGTGAGAGACACAGACGCCGGTGGCAGTGGGCTCAT	186

Query Match	8.7%	Score 433	DB 9	Length 451
Y	2356	TTCCAGGTGGATGCTTGGGAAACATCTTTCTTAAGGAACCTGAGTGGAGAACCA	2415	
Db	187	TTCCAGGTGGATGCTTGGGAAACATCTTTCTTAAGGAACCTGAGTGGAGAACCA	246	
Y	2416	TGCAGGTATCTCAGGAAGAGCTTCCTCAGGACAGGAAGATCAGAGGTGGAAGCCCTG	2475	
Db	247	TGCAGGTATCTCAGGAAGAGCTTCCTCAGGACAGGAAGATCAGAGGTGGAAGCCCTG	306	
Y	2476	GAGCCACCATTCAGTAACATCATTTGAGCATCTCTACAGCTAGGTTCCATTATGGGAA	2535	
Db	307	GAGCCACCATTCAGTAACATCATTTGAGCATCTCTACAGCTAGGTTCCATTATGGGAA	366	
Y	2536	TGGGAATATGCTGTGGGACAGAGGCTGCTGCTCCCTTCACATCTTCACACTAGGGTGG	2595	
Db	367	TGGGAATATGCTGTGGGACAGAGGCTTGGCTGTCTCCCTTTCACATCTTCACACTAGGGTGG	426	
Y	2596	TTGAGAGAGCTTTGGGAGCTAAC 2617		
Db	427	TTGAGAGAGCTTTGGGAGCTAAC 448		
RESULT 4				
AL808844		451 bp	mRNA	linear EST 19-DEC-1999
LOCUS	wf97h09.x1	Soares_NSF_F8_9W_OT_PA_P_S1	Homo sapiens	CDNA clone
DEFINITION	IMAGE:2363585 3', mRNA sequence.			
ACCESSION	AL808844			
VERSION	AL808844.1	GI:5395410		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 451)			
FEATURES	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1234 Std Error: 0.00 Seq primer: -40UP from Gldco High quality sequence stop: 444.			
FEATURES	Location/Qualifiers			
SOURCE	1. 451			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:2363585"			
	/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"			
	/lab_host="DH10B"			
	/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HR-9W pool 1: 758280-760583, 772104-774407 Soares NBHP pool 1: 304776-306311, 320136-322823, 326280-326653 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	125 a	90 c	128 g	107 t 1 others

Best Local Similarity 98.6%; Pred. No. 2.4e-46; Matches 436; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2176 TCGTCAAGTAATTTTATTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 2235
|||||
Db 10 TCGTCAAGTAATTTTATTTATGAGCGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 69
|||||

QY 2236 AGGAGAACAGCAGAGAACAGCAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 2295
|||||
Db 70 AGGAGAACAGCAGAGAACAGCAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 129
|||||

QY 2296 TCTCTAGAAATATACAGCAAGGTGAGAGACACAGACACCGGTGGAGAGGGCTCTAT 2335
|||||
Db 130 TCTCTAGAAATATACAGCAAGGTGAGAGACACAGACACCGGTGGAGAGGGCTCTAT 189
|||||

QY 2356 TTCAGAGTATGATGTGGAGACATCTTCTTAAGGAAAGCTGAGTAGGAAAGCA 2415
|||||
Db 190 TTCAGAGTATGATGTGGAGACATCTTCTTAAGGAAAGCTGAGTAGGAAAGCA 249
|||||

QY 2416 TGCAGATATCTCAGAGAAAGCTTCTCCAGGACAGAAAGATCAGAGGTGAAAGCCCTG 2475
|||||
Db 250 TGCAGATATCTCAGAGAAAGCTTCTCCAGGACAGAAAGATCAGAGGTGAAAGCCCTG 309
|||||

QY 2476 GAGCCACCATTCACATTAACATCTTACAGATCTCTACAGCTAGTCCATTATGGAA 2535
|||||
Db 310 GAGCCACCATTCACATTAACATCTTACAGATCTCTACAGCTAGTCCATTATGGAA 369
|||||

QY 2536 TGGGAATATGTGGTGGAGAGGCTGCTGCTTCCATCTTCTACACTAGGTGG 2595
|||||
Db 370 TGGGAATATGTGGTGGAGAGGCTGCTGCTTCCATCTTCTACACTAGGTGG 429
|||||

QY 2596 TTGAGAGAGCTTGGAGACTAAC 2617
|||||
Db 430 TTGAGAGAGCTTGGAGACTAAC 451
|||||

RESULT 5
A1432176
LOCUS
DEFINITION
A1432176.437 bp mRNA linear EST 30-MAR-1999
t977409.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:2114825
3' similar to contains element DBR repetitive element ;, mRNA
sequence.
A1432176
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1153 Std Error: 0.00
Seq primer: -40bp from Glbco
High quality sequence stop: 401.
Location/Qualifiers
1. 437
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2114825"
/clone_id="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; site_1: Not I;
site_2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2Nbm, pregnant uterus
NBHPU, and fetal heart NBH19) were mixed, and ss circles
were used in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT
121 a 89 c 123 g 104 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 4.4e-46; Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2176 TCGTCAAGTAATTTTATTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 2235
|||||
Db 7 TCGTCAAGTAATTTTATTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 66
|||||

QY 2236 AGGAGAACAGCAGAGAACAGCAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 2295
|||||
Db 67 AGGAGAACAGCAGAGAACAGCAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 126
|||||

QY 2296 TCTCTAGAAATATACAGCAAGGTGAGAGACACAGACACCGGTGGAGAGGGCTCTAT 2335
|||||
Db 127 TCTCTAGAAATATACAGCAAGGTGAGAGACACAGACACCGGTGGAGAGGGCTCTAT 186
|||||

QY 2356 TTCAGAGTATGATGTGGAGACATCTTCTTAAGGAAAGCTGAGTAGGAAAGCA 2415
|||||
Db 187 TTCAGAGTATGATGTGGAGACATCTTCTTAAGGAAAGCTGAGTAGGAAAGCA 246
|||||

QY 2416 TGCAGATATCTCAGAGAAAGCTTCTCCAGGACAGAAAGATCAGAGGTGAAAGCCCTG 2475
|||||
Db 247 TGCAGATATCTCAGAGAAAGCTTCTCCAGGACAGAAAGATCAGAGGTGAAAGCCCTG 306
|||||

QY 2476 GAGCCACCATTCACATTAACATCTTACAGATCTCTACAGCTAGTCCATTATGGAA 2535
|||||
Db 307 GAGCCACCATTCACATTAACATCTTACAGATCTCTACAGCTAGTCCATTATGGAA 366
|||||

QY 2536 TGGGAATATGTGGTGGAGAGGCTGCTGCTTCCATCTTCTACACTAGGTGG 2595
|||||
Db 367 TGGGAATATGTGGTGGAGAGGCTGCTGCTTCCATCTTCTACACTAGGTGG 426
|||||

QY 2596 TTGAGAGAGCT 2606
|||||
Db 427 TTGAGAGAGCT 437
|||||

RESULT 6
BE785274 921 bp mRNA linear EST 20-OCT-2000
LOCUS
DEFINITION
601474631P1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3877806 5',
mRNA sequence.
BE785274
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mhc.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DP/Genetec
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

RESULT 8
LOCUS AI432177 423 bp mRNA linear EST 30-MAR-1999
DEFINITION U97710.1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2114827
ACCESSION AI432177
VERSION AI432177.1 GI:4308441
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 423)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
Email: cgaaps-f@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 581 Std Error: 0.00
Seq primer: -400P from Glbco
High quality sequence stop: 401..
FEATURES
source
1..423
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2114827"
/clone_1lb="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site:1: Not I;
Site:2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 120 a 93 c 117 g 93 t
ORIGIN
Query Match 7.7%; Score 387.4; DB 9; Length 423;
Best Local Similarity 96.1%; Pred. No. 1.7e-40;
Matches 397; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
OY 2181 CAGTAATTTTATGAGCGCTTCTACGAGAACACAGAGAGCTTCATTCTGAGAG 2240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 11 CACTRAAACCTTATTGGGGCTTTTACGAGAACACAGAGCGCTTCATTCTGAGAG 70
OY 2241 GAAAACAGGAGAAAACAGGAGATATCTGTATTAATTTCAAGTAGTAAGTCTCT 2300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 71 GAAACAGGAGGAGAAACAGGCAATATCTGTATTAATTTCAAGTCTGTAAAGTCTCT 130
OY 2301 AGAANTATCAAGCAAGTGAAGAGACACAGACCGGCTGAGAGGGGCTCTATTTC 2360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 131 AAAATATCAAGCAGGAGGAGAGACACAAAGCACCGGTGGAGTGGGGCTCTATTTC 190
OY 2361 GATTGATGTTGGAGACATCTTTCTAAAGGAACTGAGTGGGAGGAACATGAG 2420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 191 GATTGATGTTGGAGACATCTTTCTAAAGGAACTGAGAGGAGGAACATGAG 250
OY 2421 GTATCTCAGAGAGAGCTTCTCCAGGAGAGAAATCAGAGTGAAGAGCCCTGAGCC 2480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 251 GTATCTCAGAGAGAGCTTCTCCAGGAGAGAAATCAGAGTGAAGAGCCCTGAGCC 310

OY 2481 ACCATTCAATCAATCATTTTACAGATCTCTACAGCTAGCTTCATTATGGAATGGA 2540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 311 ACCATTCAATCAATCATTTTACAGATCTCTACAGCTAGCTTCATTATGGAATGGA 370
OY 2541 ATATGTTGTGAGACAGGCTGCTGCTGCTCTCCATCTCTACACTAGGCT 2593
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 371 ATATGTTGTGAGACAGGCTGCTGCTGCTCTCCATCTCTACACTAGGCT 423
RESULT 9
LOCUS AF101960 2971 bp DNA linear GSS 06-NOV-2000
DEFINITION AF101960 Human Homo sapiens genomic clone pTMB59.14, DNA sequence.
ACCESSION AF101960
VERSION AF101960.1 GI:4193786
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2971)
Bepier, G., O'Brian, K.C., Kim, Y.C., Schreiber, G. and Pitterle, D.M.
A 1.4-Mb high-resolution physical map and contig of chromosome
segment 11p15.5 and genes in the LOH1A metastasis suppressor
region
Genomics 55 (2), 164-175 (1999)
JOURNAL 99134294
MEDLINE
COMMENT Contact: Bepier G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
Part of a 1.4 megabase contig including the LOH1A metastasis
suppressor region Bin T
Class: unknown.
FEATURES
source
1..2971
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="11p15.5"
/clone="pTMB59.14"
/clone_1lb="Human"
/note="Vector: pMOS Blue"
BASE COUNT 1121 a 589 c 514 g 747 t
ORIGIN
Query Match 5.9%; Score 296.8; DB 17; Length 2971;
Best Local Similarity 68.5%; Pred. No. 2e-29;
Matches 522; Conservative 0; Mismatches 212; Indels 28; Gaps 7;
OY 3974 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4033
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 470 TTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 529
OY 4034 AGTGGGGCAATCTTGGCTCACTGTACCTCTGCTCCAGGTCAGCGATCTCTGC 4093
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 530 AGTGCA -CAATCTTGGCTCACTGTACCTCTGCTCCAGGTCAGCGATCTCTGC 588
OY 4094 CTGACCTCTGAGTACGTGGAATATATAGGCAACACACAGCGCTGGTAATTTT 4153
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 589 CTCACCTCCCAAGATGCTGAGATTACAGGGCCGCCACACACAGCTAATTTT 645
OY 4154 TTTTCTGTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4213
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 646 -----TGTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 698
OY 4214 CCTGACCTCAAGTATCCACCGACTGCTCCCAAGTCTGGATTTACAGGTGCA 4273
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 699 TCCTGACCTC -GTGATCCACCGACTGCTCCCAAGTCTGGATTTACAGGTGCA 756
OY 4274 GCCACCATGACACGCC -ACATGTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4332
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 757 GCCACCGATCTGCGCCCAATTAATTTTATGTAAGTGAATGTAATTTTCTTAAGTGC 816

OY	4333	TTCTAAGGCGACAGTATGCACTGATGCGCTGCTTAATATCCAGACACTTTGAGGGGGCCAAAGTG	4332
Db	817	AGAAACACCCACGGCACATGGCTTCAACGCTCTAAATCCGTGGACATTTTGGGAGGCCGAGG	876
OY	4393	CGGGGATCAGCTTGAGCCCTG---GGAGTTACAGCGTGGGCAACATATGTGAGACCCCGCTC	4448
Db	877	GGCAGATACAGAGGTACAGAGATGAGAACATCCCGGCTAATACGCTAAMACCCTGCTC	936
OY	4449	TACCAAAATTT---TAAAAAATTAAGCTGGGAGTGTGGTGGCATTTGCTGTGCTCCAGCT	4504
Db	937	TACTAAAAATTAATAATTAATAATTAATTTAGCTGGGTGCAGTGGTGGGCACTGTATGCTCCAGCT	996
OY	4505	ACTTGGGAGAGCTGGAGGTGGGGAGAGCTGAAAGCTGTGAGCTGCAGGCTCCACAGGACT	4566
Db	997	ACTCGGAGGCTGAGGCGGAGAAATGGTGAACCCAGAGAGCGGAGGTTTCCAGTGAACC	1056
OY	4565	ATGATCACACACACTGCACCTTACAGCTGTAGTACAGG-----CTATTCCTCAAAAGCAACA	4618
Db	1057	AAGATCAACGCACTGCACCTTAGCCGACAGGGGAGAGCTCCGACTCAAAAAAAA	1116
OY	4619	AAATTAATGTTATCTAAACGGTAAAGGTATTAATCACAGATATATGATATGCAATTTAAAT	4678
Db	1117	AAAGATTAGAAATGAACAATGATGATATCTGCCTACCTCATTTCTCTTCAACACTGTACTA	1176
OY	4679	GAAGAAGATTAATGATTAATGATTTTAAAAATTCAAATA	4720
Db	1177	GAATTAATCTTCCAAAGGCAATCAGCGAAGAAAAAGAAAAA	1218

LOCUS	AF101960	2971 bp	DNA	1 linear	GSS 06-NOV-2000
DEFINITION	AF101960	Human Homo sapiens genomic clone pIBW59.14,			DNA sequence.
ACCESSION	AF101960				
VERSION	AF101960.1	GI:4193786			
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2971)				
AUTHORS	Bepler,G., O'Brian,K.C., Kim,Y.C., Scheider,G. and Pitterle,D.M.				
TITLE	A1.4-Mb high-resolution physical map and contig of chromosome				
	segment 1p15.5 and genes in the LOH1A metastasis suppressor				
JOURNAL	region				
	Genomics 55 (2), 164-175 (1999)				

**JOURNAL
MEDLINE
COMMENT**

Genomics 55 (2), 164-175 (1999)
99134294

Contact: Repler G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
pat of a 1.4 megabase contig including the LOH1A metastasis
suppressor region Bin T
Class: unknown.

FEATURES	SOURCE	Location/Qualifiers
		1. 2971
		/Organism="Homo sapiens"
		/db_xref="taxon:9606"
		/map="11p15.5"
		/clone="PTW859.14"
		/clone_1lb="Human"
		/note="Vector: pMOS Blue"
BASE COUNT		1121 a 589 c 514 g 74
ORIGIN		

	Query Match	5.9%	Score 293.8	DB 17	Length 2911
DB	Best Local Similarity	69.4%	Pred. No. 4,8e-29		
	Matches 489	Conservative 0	Mismatches 202	Indels 14	Gaps 6
QY	3966	CATGCTACTTTTTTTTTTTTTTTTTTTTGAGACAGGGTTTCACCTCCATCCACCAGGC	402		
DB	1138	CATGTTTCAATCTTATCTTTTTTTTTTTTTTTTGTGAGTCGGAGAGTCCTCCCGCTGCGCTAGG	1075		

QY	4026	GGAGTGACAGTGGGGGCAATCTTGGGTCACATGTAAACCTCTCCACAGGTGAAGGAT	4085
Db	1078	TGGAGTGACGT-GGCATGATCTTGGGTCTACTGCAAGCTCCGCTCTCGGGTTCACACAT	1020
QY	4086	TCCTCGCTTACGCTCCTCGTAGTGTGAATTATAGGCACACACACACGCTGGCTA	4145
Db	1019	TCCTCCGCGCTACGCTCCCGAGTAGTGTGGACTACAGGTGCCACCACTGCACCCAGCTA	960
QY	4146	ATTTTTTTTTTTTTTCTGATTTTTAGTAGAGACAGGGTTTCATCATGTGTGGCCAGGCTGG	4205
Db	959	ATTTTTTATTTT-----TATTTTAGTAGAGACGGGGTTTCCAGCTGTATCCCAAGAGTG	905
QY	4206	TCCTGAACCCCTGACCTCAAGTAGATCAACCCAGCTGGGCTCCCAAGTGTGGAGATTAC	4265
Db	904	TCCTGATCTCTGACCTC--GTGATCTGCCGCTCGGCTCCCAAGTGTGACAGATTAC	847
QY	4266	AGGTGTACGCCACCATGCACAGCCCATGTGTACATTTTTTAAATATTTTTTAAATTA	4325
Db	846	AGGGGTGACGACATGTGCTGGCTGTATTCTGATCTTGAAGAAACATTCAGTCTTACCAT	787
QY	4326	AATCT-TTATCTAAGCCAGTAGAGTAGTACCGGCTCTGATATCCACAGCATTTGAGGG	4384
Db	786	AAATATATATTTGGGCGCAGATGGGGTCTTATGCTCTGATATCCACACATTTGGGAGG	727
QY	4385	CCAAGGTCCGGGATCTAGTGGCTGG----GAGTTCAAGCTGGGCAACATAGTAGAC	4440
Db	726	CTGAGTGGGGTGATACAGAGGTCTGAGACAGCTGGGCCAATGTGAAAC	667
QY	4441	CCCGTCTCTACCAAAAATTTTAAAAAATTAAGTGGGAGTGGTGGCATTTGCTGTGGTCCC	4500
Db	666	CTCATCTCTACTATAAAAA-TACAAAAAATTAAGTGGTGGTGGGGGGGCTGTAAATCC	608
QY	4501	AGCATCTTGGGAGCTGAGTGTGGGGATGGCTCAAGCCTGTGAGTGTGAGCTGCAGTG	4560
Db	607	AGCTACTTGGGAGGCTGAGGACAGAGAAATCCTTGAACCCAGAGGGGAGAGTGGACGTG	548
QY	4561	AGCATGTATCACCACTGACATTTGAGCTGATGTACAGAGCTATCTCAAAAGCAACAA	4620
Db	547	AGCCAGAGATTTGGCACCTGACCTCAGCTGTGGGCAACAGAGTAGACACTGTCTGGAAA	488
QY	4621	ATATATGTTATCTAAAGGTAAAGGTATATATACAGAAATATATGAT	4665
Db	487	AATTAATTAATTAATTAATTAACAAAATTTATATATATATTAAT	443

RESULT	11				
LOCUS	BC033224/c				
DEFINITION	BC033224 Homo sapiens, clone IMAGE:5022014, mRNA.	1641 bp	mrna	linear	HTC 27-JUN-2002
ACCESSION	BC033224				
VERSION	BC033224.1	GI:21619936			
KEYWORDS	HTC.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REMARK	
COMMENT	
CONTACT	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk
JOURNAL	USA
TITLE	Submitted (25-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.
AUTHORS	Strausberg R
REFERENCE	1 (bases 1 to 1641) Mammalia; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Eumetazoa; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Email: cgabbs-rt@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNT)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,


```

OY 4143 CTAATTTTTTTTTTTCTGTATTTTGTAGTAGACAGAGGTTTCATCATGTTGGCCAGGC 4202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 CTAATTTT-----TGTATTTTGTAGTAGAGTGGGGTTTTCATCATGTTGGCCAGGC 162
OY 4203 TGGCTTTGAACCCCTGACCTCAAGTATCCACCCACCTCGGGCTCCCAAAGTGTGGGAT 4262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 TGGCTTTGAACCTGACCTCAAGTATCCACCCACCTCGGGCTCCCAAAGTGTGGGAT 102
OY 4263 TACAGGTGTCCAGCCACCATGCAGACCCCA 4291
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 TACAGGTGTGGGCCACCGTGTGGGCTCA 73

RESULT 13
BQ440115 814 bp mRNA linear EST 24-MAY-2002
AGENCOURT_7826937 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6152819
5', mRNA sequence.
ACCESSION BQ440115.1 GI:21179191
VERSION BQ440115
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 814)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13491 row: e column: 12
High quality sequence stop: 588.
FEATURES
Source Location/Qualifiers
1..814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6152819"
/clone_lib="NIH_MGC_67"
/tissue_type="retinodlastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: PCMV-SPORE6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 168 a 213 c 183 g 248 t 2 others
ORIGIN
Query Match 5.3%; Score 263.8; DB 14; Length 814;
Best Local Similarity 77.4%; Pred. No. 7e-25; Indels 13; Gaps 3;
Matches 362; Conservative 0; Mismatches 93;
OY 3844 AGAGATGAGTCTCACTGTGTGTCCAGAGCTGCTCGAACTCTAGGCTCAAGCAATCC 3903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 AGAGTGGGGGTCTCAATGTTGGCCAGGCTGTTTGAACCTTGGGCTCAAGCAATCC 195
OY 3904 CCGTGCCTTAGGCTCCCAAGGGGCTGGATTACAGGTGACCTACTGCACTTGACCA-- 3961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 TCGTCCCTTAGGCTCCCAAGGGGCTGGATTACAGGTGACCTACTGCACTTGACCA-- 255
OY 3962 ----ACCACTAGGTACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 AAGCATTTTCTTTCTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 315
OY 4018 ACCAGGCTGAGTGTGAGTGGGGCAATCTTGGCTCACTGTAACTCTCTCCAGGTG 4077

```

```

Db 316 GTCCAGACTGGAGTGGAGT-GGTGTGATCTTGCTGCTACTGCACCTTCACTCCAGAGTG 374
OY 4078 CAAGGATTTCTCCCTCTTACCTCTGAGTGTGATTTATTTAGGCACACACACAGC 4137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 CACATATTTCTGCTCTTATTTAGCTTCTGATGACGTGGGGCTATPAGGCATATGCACAGC 434
OY 4138 CCGGCTAATTTTTTTTTTTTCTGTATTTTGTAGTAGACAGAGGTTTCATCATGTTGGC 4197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 CCGGATTAATTTTATTTT-----TATTTTGTAGAGACGGGGTTCACCATGTTGGC 488
OY 4198 CAGGTGTGCTGAGACCCCTGACCTCAAGTATTCACACCCACCGGCTCCCAAAGTGTG 4257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 CAGGCTGTCTCAAGAGTGTGACCTCAAGTATTTTGGCCCTCCACCTCCCAAAGTGTG 548
OY 4258 GGGATTACAGGTGTGACCCACCATGCAGCCACCCACATGTTATTTT 4305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 GGGATTACAGGGGTGAGCCACTGTGCAGGCTTTTCTTTTCTTTTCTTTT 596

RESULT 14
AA921904 261 bp mRNA linear EST 23-JUN-1998
LOCUS omd0f05.s1 Soares_NFL_T.GBC.S1 Homo sapiens cDNA clone
DEFINITION IMAGE:1543521 3', mRNA sequence.
ACCESSION AA921904
VERSION AA921904.1 GI:3069213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 261)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 385 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amerisham.
FEATURES
Source Location/Qualifiers
1..261
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1543521"
/clone_lib="Soares_NFL_T.GBC.S1"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: pRTT3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH119w, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 81 a 48 c 81 g 51 t
ORIGIN
Query Match 5.2%; Score 261; DB 9; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.1e-24;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2211 GAACACAAAGAGAGCTTCATCTGTGAGGAAAGAGGCAAGCAAGCAGATATCTTG 2270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GAACACAAAGAGAGCTTCATCTGTGAGGAAAGAGGCAAGCAAGCAGATATCTTG 60

```

OY 2271 TATAATTCAGTAGATAGTATGCTGCTAGAAAATATCAGCAAGGTGAGAGACAG 2330
 DB 61 TATAATTCAGTAGATAGTATGCTGCTAGAAAATATCAGCAAGGTGAGAGACAG 120
 OY 2331 AACACCGGTGAGGTGGGCTGCTATTTCCAGTTGGATGGTGGGAACATCTTTCTTAA 2390
 DB 121 AGCACCAGGTGGGAGTGGGCTGCTATTTCCAGTTGGATGGTGGGAACATCTTTCTTAA 180
 OY 2391 GGGACCTGAGTGGGAGGAACCATGAGTATCTCAGAGAGAGCTTCTCCAGGACAG 2450
 DB 181 GGGACCTGAGTGGGAGGAACCATGAGTATCTCAGAGAGAGCTTCTCCAGGACAG 240
 OY 2451 AGATCAGCAGGTGGGAAGGC 2471
 DB 241 AGATCAGCAGGTGGGAAGGC 261

RESULT 15
 AUI48047 559 bp mRNA linear EST 05-AUG-2002
 LOCUS AUI48047 MAMMA1 Homo sapiens cDNA clone MAMMA1002454 3', mRNA
 DEFINITION sequence.
 ACCESSION AUI48047
 VERSION AUI48047.1 GI:11009568
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 559)
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Salto,K., Kawai,Y.,
 Yamamoto,J., Nakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano
 S. and Isogai,T.).
 HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
 Salto,K., Kawai,Y., Yamamoto,J., Nakamatsu,A., Ozawa,M., Nakamura
 Y., Nagai,T., Sugano,S., Isogai,T.)
 Unpublished (2000)
 JOURNAL Contact: Takao Isogai
 COMMENT Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
 Research Institute; cDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 FEATURES
 source
 1..559
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMA1002454"
 /clone_id="MAMMA1"
 /tissue_type="mammary gland"
 /note="Vector: pME18SF3"
 location/Qualifiers
 BASE COUNT 103 a 141 c 133 g 177 t 5 others
 ORIGIN
 Query Match 5.2%; Score 257.6; DB 9; Length 559;
 Best Local Similarity 77.3%; Ptd. No. 5.5e-24;
 Matches 368; Conservative 0; Mismatches 94; Indels 14; Gaps 4;
 OY 3831 GAGAGAGAGAGAGAGATGAGTCTACCTGTGTTGCCAGGCTGGTCTCAACTCTTAA 3890
 DB 8 GGGGAGAGCGGTGAGAGCAAAATCTACTGTGCTCCCAAGCTGGTCTCAAACTGCTGG 67
 OY 3891 GCTCAGAGATCCCGCTGAGTCTTACCTGCTCCCAAGGCTGGATTCAGAGTGTGACTACT 3950
 DB 68 GCTCAGAGTGTGCTGCTGCTCAGCTTCCCAAGGTGTGGATTCAGAGCATGAGCCACC 127
 OY 3951 GCACCTGACCAACCAATGATCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCA 4010

DB 128 ATGCGTGGCCCTTAA-----ATTCCTTTTGTGTTGTTTGTGTTTGAACAGCTCTCG 178
 OY 4011 CTCACATCACCAGGCTGAGTGCAGTGGGGCAATCTTGGCTCACTGTAACCTCTGCTTC 4070
 DB 179 CTCGTGCACCGAGGCTGGAGTACAGT-GGCACAAATCTCAGCTCAGCAACTCCGCCCTC 237
 OY 4071 CCAGGTGCAAGGATCTCTGCTGCTTACCTCTGAGTACCTGGAATTTATAGCACACAC 4130
 DB 238 CCGGTTCAGCAAAATCTCTGCTTCAAGCTCTCTGAGTACCTGAGTGGAGTCAAGCATGAC 297
 OY 4131 CACACGCTGCTTAATTTTCTTTTCTGATTTTATAGTATGAGACAGGGGTTTCATCA 4190
 DB 298 CACACGCTGCTTAATTTTCTTTTCTGATTTTATAGTATGAGACAGGGGTTTCATCA 355
 OY 4191 TGTGGCCAGGCTGCTTGAACCCCTGACCTCAATGATTCACACCACTGCGCTCCCA 4250
 DB 356 TGTGGCCAGGCTGCTTGAACCCCTGACCTCAATGATTCACACCACTGCGCTCCCA 413
 OY 4251 AAGTCTGGATTCAGTGTGTCAGCCACCATGACAGCCACATGATGATCTTTT 4306
 DB 414 AAGTCTGGATTCAGTGTGTCAGCCACCATGACAGCCACATGATGATCTTTT 469

Search completed: June 9, 2003, 03:29:56
 Job time : 6623.31 secs


```

Yy      4143  CTAAATTTTTTTTTTTCGTAATTTTTAGTAGAGACAGGGGTTTCATCATGTTGGCCAGGC 4202
Db      8426  CTTATGTTT-----TGATTTTTTAGTAAGTAATGGGTTTCACCTCTGTGGCCAGTC 8376

Yy      4203  TGGCTCTGAACCCCTGACCTCAAGTGAATCCACCCACCTCGGCGCTCCCAAGTGGTGGAAT 4262
Db      8375  TGGCTCTGAACCTCTGACCTCAGGTAATTCGCCACCTCGGCGCTCCCAAGTGGTGGAAT 8316

Yy      4263  TACAGGTGTCAGCCACCATGACAGCCCACTGCTACATTTTTTAAATTAATTTTTTAAT 4322
Db      8315  TACAGGTGTCAGTCACACCCCTGGCTACGTCTTAATTTTTTAAAA----- 8269

Yy      4323  TAAATGTTTTATCTAAGCCAGTAGTACGTGCTCGGCTGCTAATCCACAGACTTTTGAAG 4382
Db      8268  -----GAGGCCGGGGCGCAGTGGCTACACCCTCTGTAAATCCACAGCTTTGGGA 8223

Yy      4383  GGCCAAGGTGCGGGGATCACTTGAGCCTGGAGTTACGCGT---GGCAACATAGTGAGA 4439
Db      8222  GGTGAGATAGGTGATGATCACAAAGTTCAGAGATTGAACACAGCCTGCTGATAGTAA 8163

Yy      4440  CCGGCTCTCAACCAAAATTTAAATAATTCGTGGAGTGGTGAGCTTTGCTGTGGTCC 4499
Db      8162  ACCGGTCTCAGTAAAAA-TACAAAATTTAGCCGGGGTGGTGGCAGCAGCTGTATGTC 8104

Yy      4500  CAGCTACTTGGGAAGCTGAGGTGTGGGATGGCTGAAGACCTGTAGAGTTCGAGGCTGAGT 4559
Db      8103  CAGCTACTCAGGAGGCTGAGGCAAGAGAAATGCTTGAACCTGGGAGGGAGGAGTTGAGT 8044

Yy      4560  GAGCTATGATCACACCACTGCATCTTCAAGCTG--AGTACAGGCTATCTCAAAAGCAAC 4617
Db      8043  GAGCTGATGATGCACCACTGCATCTCCAGCCCTGGCAACAAAGCGAGACTCCCTCTCAAAA 7984

Yy      4618  AAAATATATGTTATCTAAAGGTAAGGATATATACAGAAATATATGATAGCATTTTAAT 4677
Db      7983  AAAAAAATAATATATATATATATCAAAAAATTAGCCGGCATGTAATGACGCCCTATA 7924

Yy      4678  TGAATAAGCATTAATGATTTACATGATGATTTGTAATAATCAATATACATGAAATTTCTGT 4737
Db      7923  ATCCAGGATCTCGGAGGCTGAGGACGAGAAATCGTTGAACCTAGAGAGTGAAGTTTG 7864

Yy      4738  C-TTAATATGCTAGCAACAAGCACATTTGGTTTTTACTAGAGGCACCAAGTACTTTA 4795
Db      7863  CAGTTTGTGAGATCTCCACCACTGCATCCAGCCCTGGGCAACAAGCGAGATTCACATCTC 7804

Yy      4796  AAAAAAGTTAGGCGCAGCCACACAGGGGCTCACACCTGTATATCCAGACTTTGGGAGGCCA 4855
Db      7803  AAAAAAATAAGGCCAGGTTGGTGGTGTCTACAGCCTGTATATCCAGACTTTGGGAGGCCA 7744

Yy      4856  AGGAGAGAGATCACTTGTAGGCCAGGAGTTTAGCACTGAGCAACATAGGAGATCTGA 4915
Db      7743  AGGAGAGGATGATTTGGCTGATGTACAGAGATTCGAGAAC--AGCTGGGCCAATGATGGTAA 7686

Yy      4916  TCTGTCTCTCTATAAAAAATTTAAAAAATTTGGCTAGGCCCTTTGGGCTTACACCCGTAATCC 4975
Db      7685  CCGGCTCTCTCTATAAAATATACAAAATCAGCGGGTGTCTTGGGCTGTACTTAAATCT 7626

Yy      4976  AGCACTTTGGAGGCGGAGCGGG 4999
Db      7625  AGCTTTTACGAGGCTGAGGACGG 7602

RESULT 2
US-09-764-869-1599/c
; Sequence 1599, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper

```

[illegible]

OY	4084	ATTCTCCCTGCTTAACCTCTCTGAGTAGCTGCAATATATAGCACAACACACACGACCTGGC	4143
Db	2396	ATTCCCTGCGCTCACCTCTCTGAGTAGCTGGACATACACAGGATGCGCCATGCGCCAGC	2337
OY	4144	TAAATTTTATTTTATTTTCTGTAATTTTACTAGAGACAGGGTTTCATCATGTTGGCCAGCT	4203
Db	2336	TAAATTTT-----TGATCTTGGTAGATGAGAGGTTTCACCATGTTGGCCAGCT	2286
OY	4204	GGTCTTGAACCCCTGACCTCAAGTAGATCCACCACTCTGGCCTCCCAAGTGTGAGAT	4263
Db	2285	GGTCTTGAACCTCTGACCTCAAGTAGATCCACCACTCTGGCCTCCCAAGTGTGAGAT	2226
OY	4284	ACAGGTGTACAGCCACCATGACACAGCCCAATGTACATTTTAAATTAATTTTAAT	4333
Db	2225	ATAGGATGACCACTGTGCGCCAGCCTAAATGTTAATTTT-----AATTTTAAAGTT	2171
OY	4324	AAATGTTTATCTAAGGCGAGTAGAGAGACGACCTGCGTCTGTATATCCAGCACTTTGAGG	4383
Db	2170	TTTAAATCACAAGTAGGCCACAGCGCGGTTACTCTGTATATCCCAACCTTTGGAG	2111
OY	4384	GCCAAAGTGGCGGGTCACTTAGACCTGGAGTT-----CAGCGTGGGCAACATAGGA	4437
Db	2110	GCCAAAGCAGGCGGATCACTGAGGTGAGGATTTGAGACCAACACTGGCCCAATGGGA	2051
OY	4438	GACCCCGCTCTACCAAAATTTTAAAAATTAAGCTGGAGTGTGGCATTTTCTGTGCT	4497
Db	2050	-AACCATCTTCTACTTAAAAATTAATTAAGTGTGGGCACTGAGTGCTGGCGCTGTAA	1992
OY	4498	CCGAGCTACTTGGGAGCTGAGGTGTGGGAGTGTGAGGCTGAGGCTGTGAGGCTGAG	4557
Db	1991	CCGAGCTACTTGGGAGGCTGTGGGAGGATTCGTTGACATGAGAGCGGAGGCTGGA	1932
OY	4558	GTGAGCTATGATCAACACCATTGCACTTCAAGCCTAGTAGACAGGCTATCTCAAAAGCAAC	4617
Db	1931	GTGAAGTAGATGTGCGCACCTGCATCGACCTGAGTTGGGGA-----TAAGAGCAAACTCTG	1877
OY	4618	AAATTAATGTTTATCTAAAGGTAAGGTATATACAGAAATATATGATAGCACTTTAAT	4677
Db	1876	TCTCAAAAAAATAAATAATTTCTATTATGAAAAAAGAAATTTAAACCAACACACTTT	1817
OY	4678	TGAATAAGCATTAATGATATACATGATGTGTAATATCAATATACATGAATCTTGCT	4737
Db	1816	ACATGTGATCATTTCCCTCACTCACTCAAAAAAAGGTTTACTGTGATG	1757
OY	4738	CTTAATAATGCTAGCAACAMGCACTTTGGTTTACTAGGCGCACCAAGTACTTTAAA	4797
Db	1756	CTAAGTGTGCT-----CTAAGAGTTAGAAAT	1730
OY	4798	AAAAAGTTAAGGCCACGACAGGGCTACACCTGTAATCCAGACATTTGGAGGCCAG	4857
Db	1729	TCCAAAAATAGCCACAGACACTGTGCTCACTGTAATTCAGACACTTTGGAGGCCATG	1670
OY	4858	GCAGAGAGATCACTTGAAGCCAGAGATTTAGGACTAGCAACATAGGAGATCTCATC	4917
Db	1669	ATGGGTGATCACTTGAAGCCAAAGTTCGAGATGTGTGGCCAA--CAAGTAGAAC	1612
OY	4918	TTGTCTCTATAAAAATTAATAATTTGGCTTAGGCTTTGGCTTACACCGGTATCCAG	4977
Db	1611	CTGTCTCTACTGAATAATACAAAAATTAAGTGGGATGTGGCGGAGGCTGGCAATCCAG	1552
OY	4978	CACTTTGGGAGCGCGAGGC	4996
Db	1551	CTACTCGGAGTAGTGAAGC	1533

RESUDA 5
US-10-161-803-34
; Sequence 34, Application US/10161803
; Publication No. US2003092028A1
; GENERAL INFORMATION:
; APPLICANT: Ma, Yuanhong
; APPLICANT: Lih, Chih-Jian

```

: APPLICANT: Chen, Fan
: APPLICANT: Faltman, Jeffery
: APPLICANT: Chen, Yil-Der I.
: TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS AND
: TITLE OF INVENTION: TREATMENT OF INSULIN RESISTANCE AND RELATED CONDITIONS
: FILE REFERENCE: 421452000300
: CURRENT APPLICATION NUMBER: US/10/161,803
: CURRENT FILING DATE: 2002-06-03
: PRIOR APPLICATION NUMBER: US 60/295,264
: PRIOR FILING DATE: 2001-06-01
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 34
: LENGTH: 4360
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-161-803-34

```

Query Match	6.7%	Score	336.6	DB	9	Length	4360
Best Local Similarity	71.2%	Pred. No.	5.3e-83				
Matches	546	Conservative	0	Mismatches	179	Indels	42
						Gaps	6

QY	384	AGAGATGAGAGCTACAGCTGTGTGTGTGCACAGCGTGGCTGCAGAACTCCTAGGCTCAAGCAATCC	3903
Db	1240	AGAAACAGAGGGTTTTTACCCTATTTGGCGAGGGGTCTTGAATCCTCGACTGGGGTGAATCT	1299
QY	3904	CCCTGCACTTACCTCCCAAGAGGGGTGGGATTACAGAGTGTAGCTACTGACTTG-----	3957
Db	1300	ACCCTCCTTGGCCCTCCCAAAGTGGTAGATTACAGGCAATGAGCCACGCTGCTGGCCCTA	1355
QY	3958	---ACCAACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGGGTTTCATCC	4014
Db	1360	ATTACCCAAATCTTTTTTGT	1419
QY	4015	ATCACCCAGAGTGGAGTGCAGTGGGGGCAATCTTGGCTCACTGTATACCTCTGCTCCAG	4074
Db	1420	GTCCGCCCAAGGTGGAGTGCAGT-66CGTGAATCTTGGCTCACTGTGAAAGTCTCGCTCCCG	1478
QY	4075	GTCGACAGGATTTCTCGCTTACCTCCTAGTAGAGTGGAAATTATAGGCACACACCAC	4133
Db	1479	GTTACAGCGAATTTCTCTGCTCCTCAGCTTCCGAGTAGCTGGGAACTACAGGCCCCCAACC	1533
QY	4135	ACGCGCTGGCAATTTTTTTTTTTTTTTTTTTCTGTATTTTGTAGTAGACAGGGTTTCATCATGTT	4194
Db	1539	ACGCCAGATTAA-----TTTGTGTATTTTGTAGTAGAGAGCGGTTTTACCGTGT	1588
QY	4195	GGCCAGCGCTGTCTGTAACCCCTGACCTCAAGTATCCACCACCTCGGCTCCCAAGT	4255
Db	1590	AGCCAGGATGTCTCGAATCTTCTGACCTCA--TGATCGCCACACTCGCTCCCAAGT	1647
QY	4255	GCTGGGAATTACAGGTGCAGCGCCACCAATGACAGCCCAATGTACATTTTTTAAATATAT	4314
Db	1648	GCTGGGAATTACAGGTGCAGCGCCACCGTCCCGGCCCTA-----AT	1688
QY	4315	TTTTTAATTAATAATGTTATCTAAGGCGAGTAGAGTACGCGCTGTATCCACAGA	4374
Db	1689	TACCACTTCTTAAGAAAAAGTGGCGCGGCGGCGAGTGGCTGAACCTGTATATCCACAGA	1748
QY	4375	CTTTGAGGGGCCAAGTCCGGGGATCA--TTGAGCTGGGAGTTCAAGCGTGGCCACAT	4432
Db	1749	CTTTGGGAGGCGCGAGGAGGTGATCAAGTACAGAGATGAGACATCTCGGCTAACAC	1808
QY	4433	AGTAGAGACCCCGTCTCAACAAAAATTTAAAAAATTAGCTGGAGTGTGGCATTTGCT	4492
Db	1809	GGTGAACCAACGCTCTCAATAAAAAATCAAAAAATTTAGCGGGGTGTGGCGGGCCCT	1868
QY	4493	GTCGCTCCCACTACTTTGGGAGCTGAGAGTGTGGGATGAGCTGGAAGCTGTGAGGTCGAGG	4552
Db	1869	GTATCTCCCACTACTCGGAGGCGGAGGCGAGAGCAATGGCCTGAACCTGGGAGGCAAGC	1928
QY	4553	CTGCAAGTAGCTATGATCAACCACTGCACTTCAAGCTGAGTGAACAG	4599
Db	1929	TTGCAAGTAGCTATGATCAACCACTGCACTTCAAGCTGAGTGAACAG	1975

TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-7028

Query Match
Best Local Similarity 61.7%; Score 331.6; DB 9; Length 32194;
Matches 719; Conservative 0; Mismatches 374; Indels 73; Gaps 9;

3845 GAGATGAGATCTCAGCTGTGTGTCCAGGCTGCTGTGAACTCTAGGCTCAACATCC 3904
11268 GAGACGGGGTGTGGCCACATGCTGTGAACTCTGTGAACTCTAGGCTCAACATCC 11209
3905 CCGGCTTACCTCCCAAGGGGCTGGATTACAGTGTGAGTCTGACACTTACCAACC 3964
11208 CCGGCTTACCTCCCAAGGGTGTGGATTATATGCTGTGAGTCTGACACTTACCAACC 11153
3965 ACATGCTACTTGT 4024
11152 TGT 11101
4025 CTGAGTGTGAGTGGGGCAATCTTGGCTCTACTCTAATCTGTGCTTCCAGGTCAAGCA 4084
11100 CAGAGGAGCAGTGGCAATCTTGGCTCTACTCTAATCTGTGCTTCCAGGTCAAGCA 11041
4085 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4144
11040 TCTCTC-ACCTCTCTCTCTCTGAGCAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCT 10982
4145 AATTT 4204
10981 AATTT 10931
4205 GTCTTGAACCCCTGACCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4264
10930 GT-----GTCAAGCAATCTCTCACTGTGGCTCTCAAGTGTGAGTGTGAGTGTG 10885
4265 CAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4324
10884 CAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10847
4325 AATTT 4384
10846 AATTT 10787
4385 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4438
10786 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10727
4439 ACCGCTGTCTACCAAAATTTTAAATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4498
10726 ACCGCTGTCTACCAAAATTTTAAATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10667
4499 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4558
10666 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10607
4559 TGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4613
10606 TGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10547
4614 AATCAAAATATATGTTTATCTAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4673
10546 AATCAAAATATATGTTTATCTAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 10487
4674 AATTTGAAAAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 4733
10486 CACTTTGAAAGCCAGGTGGAGATGCT--TAAAGCCAGGATTCAGAGCTGTGAGC 10429
4734 TGTCTTAAATATATGCTGAGCAAGGCAATTTGTTTACTAGGCAACCAAGTACT 4793
10428 AGCATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10369
4794 TAAATAAATGTTAGGCGAGCAGGAGGCTCACAACCTGTATCCAGCACTTTGGAGGC 4853

10368 AATTAATTAATATAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10309
4854 CAGAGGAGAGATCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4913
10308 CAGAGGAGAGATCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10251
4914 GATCTGTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4973
10250 ACCGCTGTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10191
4974 CAGCACTTTGGAGGCGGAGCGGG 4999
10190 CAGCACTTTGGAGGCGGAGCGGG 10165

RESULT 12
US-10-091-483-335/c
Sequence 335, Application US/10091483
Publication No. US20030049650A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT12C1
CURRENT APPLICATION NUMBER: US/10/091,483
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ. ID NOS: 348
Prior Application removed - See File Wrapper or Palm
SOFTWARE: Patentln Ver. 2.0
SEQ. ID NO 335
LENGTH: 10139
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-483-335

Query Match
Best Local Similarity 69.6%; Score 329; DB 9; Length 10139;
Matches 560; Conservative 0; Mismatches 210; Indels 35; Gaps 7;

3838 AGAGAGAGATGAGTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3897
3855 AGAGAGAGATGAGTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3796
3898 CAATCCCCCTGCTTACCTCCCAAGGGGCTGGATTACAGTGTGAGTGTGAGTGTGAGTGTG 3957
3795 TGAATCTCCACACCTCAGCTACCAAACTGTAAATTAAGGATGAGGATGAGGATGAGG 3740
3958 ACCAACCAGATGATCTTT 4017
3739 TCCAGTCTCAATATATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3680
4018 ACCAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4077
3679 TGCCAGGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3620
4078 CAAGCATTTCTCTGCTTACCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4137
3619 CAGCATTTCTCTGCTTACCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3560
4138 CTTGCTTAATTT 4197
3559 CCGGCTTAA-----TTTTGTGAATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 3509
4198 CAGGCTGTCTTGAACCCCTGAGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4257
3508 CAGGCTGTCTTGAACCCCTGAGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3450
4258 GGGATTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4317
3449 GGGATTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3399
4318 TTAATTAATATGTTTATCTAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4377

Db 3398 AACCTAAATATGT-----GGGCCAGCGCCAGTGGCTACGCGCTGTAATCCCAACACTT 3344
QY 4378 TGAGGGGCGAAGGTGGGGGATCATCTGAGCCCTGGAGTT-----CAGCGTGGGACACA 4431
Db 3343 TGGGAGGCGGAGGCGGAGATCATCTGAGGTGAGGATTCAGACAGCGCTGGCCACA 3284
QY 4432 TAGTGAGCCCGCTCTCTACCAAAATTTAAAAATTTAGCTGGAGTGGTGGCATTTGCC 4491
Db 3283 CAGTAAACCCCGTCTCTA-CAAAAATAATCAAAATTAAGTGGGATGATGGCAGGTGCC 3225
QY 4492 TGTGTCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAGCGTGGAG 4551
Db 3224 TGTATCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAGCGTGGAG 3165
QY 4552 GCTGAGTGGAGTATGATCACAACCACTGCACTTCACTGAGTGGAGTGGCTATCTCAAA 4611
Db 3164 GTTGAGTGGAGCGGAGATCAGCCACTGCACTCCAACTGGGAGACAGCAAGACTAG 3105
QY 4612 GCAACAAATATATGTTTATCTAAA 4636
Db 3104 TCTCTCTCAAAAGATTTAAAAAAA 3080

RESULT 13

US-09-764-846-335/c
; Sequence 335, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic acids, proteins, and antibodies
; FILE REFERENCE: PT12
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 335
; LENGTH: 10139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-846-335

Query Match 6.6%; Score 329; DB 10; Length 10139;
Best Local Similarity 69.6%; Pred. No. 1.3e-80;
Matches 560; Conservative 0; Mismatches 210; Indels 35; Gaps 7;
QY 3838 AGAGAGAGATGAGATCTCACTGTGTGTCAGGCGTGTCTGCAACTCTAGGCTCAAG 3897
Db 3855 AGAGACAGAGTGTTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 3796
QY 3898 CAATCCCTGCTTACGCTCCCAAGGGCTGGGATTAACAGTGTGAGTCTACTGCACTTG 3957
Db 3795 TGATCTGCCACACGACCTACCAAACTGTTAAGATTAACAGCATGAGGCACTG---TG 3740
QY 3958 ACCAACCAGATGATTTTGT 4017
Db 3739 TCCAGTCTCATATATGTTGT 3680
QY 4018 ACCAGGCTGAGTGCAGTGGGGCAATCTGGCTCACTGTAACTGTGCTGCCCTCCAGGTG 4077
Db 3679 TGGCAGGTTGAGTGCAGTGGGCAATCTGGCTCACTGTAACTGTGCTGCCCTCCAGGTG 3620
QY 4078 CAGGAGTTCCTGCTTACGCTCCCAAGGGCTGGGATTAACAGTGTGAGTCTACTGCACTTG 4137
Db 3619 CAACCAATTCCTGCTTACGCTCCCAAGGGCTGGGATTAACAGTGTGAGTCTACTGCACTTG 3560
QY 4138 CCGGCTAATTTTGT 4197
Db 3559 CCGGCTAATTTTGT 3509
QY 4198 CAGGCTGTCTGAACCCCTGACCTCAAGTATCCAGCCAGCTGGGCTCCCAAGGTCT 4257
Db 3508 CAGGCTGTCTGAACCCCTGACCTCAAGTATCCAGCCAGCTGGGCTCCCAAGGTCT 3450

QY 4258 GGGATTACAGGTGTCAAGCCACCAATGACAGCCACATGATGATTTTAAAAATTTATTT 4317
Db 3449 GGGATTACAGGTGTCAAGCCACCAATGACAGCCACATGATGATTTTAAAAATTTATTT 3399
QY 4318 TTAATTAATATGTTTATCTAAGCCAGTACAGTACGCTGCTGTATATCCAGCACTT 4377
Db 3398 AACCTAAATATGT-----GGGCCAGCGCCAGTGGCTACGCGCTGTAATCCCAACTT 3344
QY 4378 TGAGGGGCGAAGGTGGGGGATCATCTGAGCCCTGGAGTT-----CAGCGTGGGACACA 4431
Db 3343 TGGGAGGCGGAGGCGGAGATCATCTGAGGTGAGGATTCAGACAGCGCTGGCCACA 3284
QY 4432 TAGTGAGCCCGCTCTCTACCAAAATTTAAAAATTTAGCTGGAGTGGTGGCATTTGCC 4491
Db 3283 CAGTAAACCCCGTCTCTA-CAAAAATAATCAAAATTAAGTGGGATGATGGCAGGTGCC 3225
QY 4492 TGTGTCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAGCGTGGAG 4551
Db 3224 TGTATCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAGCGTGGAG 3165
QY 4552 GCTGAGTGGAGTATGATCACAACCACTGCACTTCACTGAGTGGAGTGGCTATCTCAAA 4611
Db 3164 GTTGAGTGGAGCGGAGATCAGCCACTGCACTCCAACTGGGAGACAGCAAGACTAG 3105
QY 4612 GCAACAAATATATGTTTATCTAAA 4636
Db 3104 TCTCTCTCAAAAGATTTAAAAAAA 3080

RESULT 14

US-09-860-670-250/c
; Sequence 250, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruden et al.
; TITLE OF INVENTION: Nucleic acids, proteins, and antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 250
; LENGTH: 14448
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-250

Query Match 6.6%; Score 328; DB 9; Length 14448;
Best Local Similarity 69.8%; Pred. No. 3.2e-80;
Matches 561; Conservative 0; Mismatches 175; Indels 68; Gaps 6;
QY 3844 AGAGATGAGTCTCACTGTGTGTCAGGCTGTGTCGAATCTTAGGCTCAAGCAATCC 3903
Db 9307 AAAGATGGGTCTCATATGTTGTCAGGCGCGGCTTGAATCTTGCCCTCAAGTATCC 9248
QY 3904 CCGTGCCTTACGCTCCCAAGGGGCTGGGATTAACAGTGTGAGTCTACTGCACTTbACCAAC 3963
Db 9247 TCTGCTGTGGCTCTCAAGGT 9191
QY 3964 CACATGTACTTTTGT 4022
Db 9190 -----CGAATTAATTTTAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9136
QY 4023 GCGTGAAGTCAAGTGGGGCAATCTGGCTCACTGTAACCTGTGCTCCAGGTGCAAGC 4082
Db 9135 GCGTGAAGTCAAGT-GGCAAGATCTGCGCTGACCCCAACCTCACTCCCGGTTCAAGC 9077
QY 4083 GATTCCTGCGCTTGGCTCCGAGTACGTGAATTTATAGGACACACCAAGCGCTGG 4142
Db 9076 GATTCCTGCGCTTGGCTCCGAGTACGTGAATTTATAGGACATGACACCAACCA 9017


```

QY 4143 CTAATTTTTTTTTTTCTGATATTTTATAGACAGAGGTTTCATCATGTTGGCCAGGC 4202
    |||||
Db 9016 ATAAATT-----TCATATTTTAGTAGACAGAGGTTTTCACATTTGGTCAAGC 8968
QY 4203 TGTGCTTGAACCCCTGACCTCAGTAGTATCCACCACTCGGCTCCCAAAAGTGTGGAT 4262
    |||||
Db 8967 TGTGCTTGAACCCCTGACCTCAGTAGTATCCCTCCACCTCGGCTCCCAAAAGTGTGGAT 8908
QY 4263 TACAGGTGTGACGACCATGACATGATGATATTTTAAATTTTAAATTTTAA- 4321
    |||||
Db 8907 TACAGGATGTGACGACCATGACATGATGATATTTTAAATTTTAAATTTTAAAGAG 8848
QY 4322 -----TTAAATGTTATCTAAAGC 4341
    |||||
Db 8847 GACCCAGAGATCAAAAGTGCCTATTAACCCAGAAAGTCTAAATGAGACAAAGAGAGC 8788
QY 4342 CAGTAGCAGTAGCTCGGCTGTATATCCAGCACTTTGAGGGGCCAAAGTGGCGGATCA 4401
    |||||
Db 8787 TGGGCGTGTGGCTCAGTCTCTGTATATCCAGCACTTTGAGGGGCCAAAGTGGCGGATCA 8728
QY 4402 CTGAGCCTGGAGTT-----CAGCGTGGGCAACATAGTAGACCCCGCTCTAACCAA 4455
    |||||
Db 8727 CCGAGGTGAGAGATTCAGACACAGCCTGGCCACATATGTAACCTGTCTCTACTAA 8668
QY 4456 AATTTAAAAAATTAGCTGGAGTGTGGCAATTTGCTGTGCTCCACACTACTTGGGAAGC 4515
    |||||
Db 8667 AAACACAAAAAATTAGCTGGAGTGTGGCAATTTGCTGTGCTCCACACTACTTGGGAAGC 8608
QY 4516 TGAAGTGTGGGATGTGCTGAAGCCTGTGAGGTGAGAGTGCATGAGCTATGATCACACC 4575
    |||||
Db 8607 TGAAGGAGAGAAATTCCTGTGAACCTAGAGGACAGAGCTCAGTGGCTGAGATCGACAC 8548
QY 4576 ACTGCACTTCAGCCTGAGTGACAG 4599
    |||||
Db 8547 ACTGCACTTCAGCCTGAGTGACAG 8524

RESULT 15
US-09-860-670-253/c
; Sequence 253, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; PRIOR FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 253
; LENGTH: 14451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-253

```

Query Match 6.6%; Score 328; DB 9; Length 14451;
 Best Local Similarity 69.8%; Pred. No. 3.2e-80;
 Matches 561; Conservative 0; Mismatches 175; Indels 68; Gaps 6;

```

QY 3844 AAGATGAGAGTCTCACTGTGTTGTCCAGGCTGCTCGAACTCTAGGCTCAAGCAATCC 3903
    |||||
Db 9307 AAGATGAGAGTCTCACTGTGTTGTCCAGGCTGCTCGAACTCTAGGCTCAAGCAATCC 9248
QY 3904 CCTGCTTACGCTCCAGAGGGGCTGGATTAACAGTGTGAGTACTGCACTTGACCAAC 3963
    |||||
Db 9247 TCTGCTTACGCTCCAGAGGGGCTGGATTAACAGTGTGAGTACTGCACTTGACCAAC 9191
QY 3964 CACATGCTACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4022
    |||||
Db 9190 -----CGAATATATATTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9136
QY 4023 GGCTGAGTGTGAGTGGGGCAATCTTGCTCACTGTAACTCTGCTCCAGGTCAGAGC 4082

```

```

Db 9135 GGCTGAGTGTGAGTGGGGCAATCTTGCTCACTGTAACTCTGCTCCAGGTCAGAGC 9077
    |||||
QY 4083 GATTCCTGCTTACGCTCCAGAGGGGCTGGATTAACAGTGTGAGTACTGCACTTGACCAAC 4142
    |||||
Db 9076 GATTCCTGCTTACGCTCCAGAGGGGCTGGATTAACAGTGTGAGTACTGCACTTGACCAAC 9017
    |||||
QY 4143 CTAATTTTTTTTTTTCTGATATTTTATAGACAGAGGTTTCAATGTTGGCCAGGC 4202
    |||||
Db 9016 ATAAATT-----TCATATTTTAGTAGACAGAGGTTTTCACATTTGGTCAAGC 8968
QY 4203 TGTGCTTGAACCCCTGACCTCAGTAGTATCCACCACTCGGCTCCCAAAAGTGTGGAT 4262
    |||||
Db 8967 TGTGCTTGAACCCCTGACCTCAGTAGTATCCCTCCACCTCGGCTCCCAAAAGTGTGGAT 8908
QY 4263 TACAGGTGTGACGACCATGACATGATGATATTTTAAATTTTAAATTTTAA- 4321
    |||||
Db 8907 TACAGGATGTGACGACCATGACATGATGATATTTTAAATTTTAAATTTTAAAGAG 8848
QY 4322 -----TTAAATGTTATCTAAAGC 4341
    |||||
Db 8847 GACCCAGAGATCAAAAGTGCCTATTAACCCAGAAAGTCTAAATGAGACAAAGAGAGC 8788
QY 4342 CAGTAGCAGTAGCTCGGCTGTATATCCAGCACTTTGAGGGGCCAAAGTGGCGGATCA 4401
    |||||
Db 8787 TGGGCGTGTGGCTCAGTCTCTGTATATCCAGCACTTTGAGGGGCCAAAGTGGCGGATCA 8728
QY 4402 CTGAGCCTGGAGTT-----CAGCGTGGGCAACATAGTAGACCCCGCTCTAACCAA 4455
    |||||
Db 8727 CCGAGGTGAGAGATTCAGACACAGCCTGGCCACATATGTAACCTGTCTCTACTAA 8668
QY 4456 AATTTAAAAAATTAGCTGGAGTGTGGCAATTTGCTGTGCTCCACACTACTTGGGAAGC 4515
    |||||
Db 8667 AAACACAAAAAATTAGCTGGAGTGTGGCAATTTGCTGTGCTCCACACTACTTGGGAAGC 8608
QY 4516 TGAAGTGTGGGATGTGCTGAAGCCTGTGAGGTGAGAGTGCATGAGCTATGATCACACC 4575
    |||||
Db 8607 TGAAGGAGAGAAATTCCTGTGAACCTAGAGGACAGAGCTCAGTGGCTGAGATCGACAC 8548
QY 4576 ACTGCACTTCAGCCTGAGTGACAG 4599
    |||||
Db 8547 ACTGCACTTCAGCCTGAGTGACAG 8524

```

Search completed: June 9, 2003, 05:30:55
 Job time : 678.38 secs

OY	4013	CCATCACCCAGGGCGAGAGTGCAGTGGGGGGAACCTTGGCTACATGTAACTCTGCTCCC	4072
Db	13535	CTGTCCACCGAGCGGTGAGTACAGT--GGCACTATCTCACTCACTGGAACCTCTGCTCCT	13583
OY	4073	AGGTGCAGGCAATCTCTCCTCCTTAGCGCTCTGAGTAGTGAATATATAGACACACCA	4132
Db	13584	GGGTTCAAAGGAATCTGTGCTCTCAGCTCTCTAGTAGTCTGGGATTTACAGCATGACCA	13643
OY	4133	CCACGCTGGCTAATTTTTTTTTTTTTTCTGTATTTTTTAGTAGACAGGGTTTCATCAT	4192
Db	13644	CCATGCCCCCGTTAATTTTTT-----TGTATTTTAGTAGAGAC--GGTTTCAACGGG	13694
OY	4193	TTGGCCAGGGCGGTCTGAACCCCTGACCCATGATATCAACCCACTGGCTCCCAAA	4252
Db	13695	TTGCCACGATGTCTTGAACCTCTGACTCAAGTAGTTACCCACTCAGCTCCCAAA	13754
OY	4253	GTGCTGGGATTACAGGTGTGACGCCACCATGACAGCCACATGTATTTTTTAAATTT	4312
Db	13755	GTGCTGGGATTATATATGTGTAGCGACATCGGCTA-----	13788
OY	4313	ATTTTAAATTTAAATTTTATCTAAAGCCAGTACAGTACTGCGCTGTATCCAG	4372
Db	13789	--TGTATTTTAAAAAAGTGGGTCTATGGGGCTGGGGCTGGGTCTCATGCTGTATCCAG	13846
OY	4373	CACATTAAGGGGCGCAAGTGTGGGGGATCACTTAGGCTG---GAGTTTCAAGCTGGGCA	4428
Db	13847	CACATTTGTGACACCGAGCGGGGTGATCACAAGTCAAGGATCGAGACATCTGCTTA	13906
OY	4429	ACATATGTAGACCCCGCTCTCTACCAAAATTTTAAAAATTTAGCTGGAGATGTGGCATTT	4488
Db	13907	ACAGGGTGAACCCCGCTCTACTATAAAATACAAAAATTAACCAAGCATGTGGTGGGC	13966
OY	4489	GCCGTGTGTCCTCCAGCTACTTGGGAAGCTGAGGTGGGGATGGCTGAAGCTGTGAGGTC	4548
Db	13967	GCCGTGTGTCCTCCAGCTACTGCGGAGGCTGAGCGCAGGAATGGCTGGAACCTGGGAGCGG	14026
OY	4549	GAGGCTCAGTGAAGCTATGATCACACACACTGACCTCAGCGTAGAGACAG	4599
Db	14027	GAGCTTCAATGAGCGGAGATCACGCCACCGTACTCCAGCTGAGGACAG	14077
RESULT 2			
US-08-367-841A-43			
: Sequence 43. Application US/08367841A			
: Patient No. 6319687			
GENERAL INFORMATION:			
APPLICANT: Chader, Gerald J.; Rodriguez,			
APPLICANT: Ignacio R.; Mazuruk, Krzysztof;			
APPLICANT: Tombran-Tink, Joyce			
TITLE OF INVENTION: PIGMENT EPITHELIUM			
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC			
NUMBER OF SEQUENCES: 43			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Morgan & Flanagan			
STREET: 345 Park Avenue			
CITY: New York			
STATE: New York			
COUNTRY: USA			
ZIP: 10154			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC Compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: WORDPERFECT 5.1			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/367,841A			
FILING DATE: 30-DEC-1994			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/257,963			
FILING DATE: 07-JUN-1994			

[illegible]

Db 2968 TTATCTAAAAAGTCTTTTAAAGCGCGGCGGCTGCTCAGCCCTGTATATCCACACACTT 3027
OY 4378 TGAGGGGCGCAAGTGGCGGGGATCACTTGAAGCTG----GAGTTGAGCGGTGGGCAACATA 4433
Db 3028 TGGGAGGCGGAGGGGGGCGATCAAGAGTCAAGAGATCGAGACATCCCGCTAAACG 3087
OY 4434 GTGAGACCCCTCTCTACCAAAATTTAAAAATTAGTGGGAGTGGTGGCATTTGGCTG 4493
Db 3088 GTGAACCCCTCTCTACTATAAAATACAAAAATTAGCGGCGGTAGTGGCGGCGCTG 3147
OY 4494 TGTCCAGCTACTTGGGAGCTGAGTGTGGGAGTGGCTGAAGCTGTGAGGTGAGGC 4553
Db 3148 TAGTCCAGCTACTTGGGAGCTGAGGAGAGAGATGGCTGAACCGGAGGCGGAGCT 3207
OY 4554 TGCAGTAGCTATGATCAACACACTGCACTTCAAGCTGAGTGAAGTCAAGCTCAAAAGC 4613
Db 3208 TGCAGTAGCGGAGTCCCGGCACTGCACTCCAGCTGGGGGAGAGAGACACTCCGTC 3267
OY 4614 AAACAAATAA 4624
Db 3268 TCAAAAAAAA 3278

RESULT 6
US-08-724-394A-21
Sequence 21, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchinashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereeto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3634
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Plets, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21
Query Match 6.5%; Score 327.4; DB 2; Length 246240;
Best Local Similarity 72.9%; Pred. No. 6,7e-78;
Matches 489; Conservative 0; Mismatches 151; Indels 31; Gaps 4;

OY 3974 TTTTCTTTTTTTTTTTTTTTTTTTTGTGAGACAGGTTTCATCTCATCAACCGGTGAGTGC 4033
Db 2619 TTTTATTTTATTTTATTTTATTTTGTGAGACGTTGCTCACCCTGCGCCAGGCTGAGATGC 2678
OY 4034 AGTGGGGCAATCTGGGCTCACTGTACCTCTGCTCCAGGTGACAGGATTCCTCTGC 4093
Db 2679 AGT-GGTGAGTCTTGGCTCACTGTACCTCCAGCTCCAGGTTCAAAGCATTCCTCTGC 2737
OY 4094 CTTAGCTCTAGTAGTGTGATTAATATAGACACACACACACAGCTGCTGCTAATTTTTTTT 4153
Db 2738 CTCATCTTCCCAAGCAGCTGGGATTAACAGTGTGCTGTGCCACACATGCCAGCTAATTTT 2794
OY 4154 TTTTCTGTATTTTATTTAGTAGACAGGTTTCATCATCTTGGCCAGGCTGTCTGAAC 4213
Db 2795 -----TGTATTTTATTTAGTAGACAGGTTTTCGCGTGTGGCCAGGCTGTCTGAAC 2847
OY 4214 CCGTACCTCAAGTATGATCCACCCAGCTGCGCTCCCAAGTGTGGATTACAGTGTCA 4273
Db 2848 TCTTACCTCAAGTATGATCCCGCTGCGCTCCCAAGTGTGGATTACAGTGTCA 2907
OY 4274 GCCACATGACACAGCCCAATGTTAATTTT-----TTTAAATTTATTTT 4317
Db 2908 GCCACTGGCGCCAGCGGGGTTTCATCTTAATATATATATATATATATATATATATATAT 2967
OY 4318 TTAATTAATTTTATATCAAGGCAATGACGATGCTGCTGTATATCCAGACTT 4377
Db 2968 TTATCTAAAAAGTCTTTTAAAGCGGGGCGGTGCTACCCCTGTAAATCCAGACTT 3027
OY 4378 TGAGGGGCGCAAGTGGCGGAGTCACTTGAAGCTG----GAGTTGAGCGGTGGGCAACATA 4433
Db 3028 TGGGAGGCGGAGCGGGGAGTCAAGAGTCAAGAGATGAAGATATCCCGCTTAAAGG 3087
OY 4434 GTGAGACCCCGTCTCTACCAAAATTTAAAAATTAGTGGGAGTGGGCAATTTGCCG 4493
Db 3088 GTGAACCCCGTCTCTACTATAAAATACAAAAATTAGCGGCGGTAGTGGCGGCGCTG 3147
OY 4494 TGTCCAGCTACTTGGGAGTGAAGTGTGGGATGCTGAAGCCTGTGAGTGCAGGC 4553
Db 3148 TAGTCCAGCTACTTGGGAGTGAAGTGTGGGAGTGAAGTGAAGCCTGGAGGCGGAGCT 3207
OY 4554 TGCAGTAGCTATGATCAACACACTGCACTTGAAGCTGAGTGAAGGCTATCTCAAAAGC 4613
Db 3208 TGCAGTAGCGGAGTCCCGGCACTGCACTCCAGCTGGGGGAGAGAGAGACTCCGTC 3267
OY 4614 AAACAAATAA 4624
Db 3268 TCAAAAAAAA 3278

RESULT 7
US-08-724-394A-22
Sequence 22, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchinashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereeto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor

OY		3898	CATCCCCCTGCTTACCTCCCAAGGGCTGGATTACAGGTGAGCTACTGCACCTTG	3957
Db		4351	TAACTCTTCACC-TAGTCTCCGGGTAGCTGCACAACCAAAG-----	4391
OY		3958	ACCAACACATAGTACTTTTTTTTTTTTTTTTTTTTTTTTGTGACACAGGGTTTCACTCCATC	4017
Db		4392	-----CACCGGTTTATCTCATCTCTTTTTTTTTTTCTTTGAGACTGAGTCTTCTCTGA	4446
OY		4018	ACCCAGGCTGGAGTGCAGTGGGGCAATCTTGGCTACATGTAACTCTCTGCCCTCCACAGTG	4077
Db		4447	GCCAGGCTGGAGGGCAGT-GGGGTGATCCGGCTACATGCACAATCCGCTTCAAGGTTT	4505
OY		4078	CACACGATTCCTCCGCTTAGCCCTCCGATTAAGTACTGGAATTAATAGGCACACACACACAG	4137
Db		4506	CACACAAATCTCTCTGCTCAGCCTCTGAGATGGCTGGGAACTACAGGGGTGGCACCAGA	4565
OY		4138	CCTGGCTAATTTTTTTTTTTTTTTTTTTCTGTATTTTTTATAGTAGACAGGGTTTATATATGTGGC	4197
Db		4566	GCGAGTTAATTTTTTTTTTTTTTTTTTTGTATTTTTTATGTGACACTGGTTTCACTATATGGT	4625
OY		4198	CAGGCTGCTCTTGAACCCCTGACCTCAAGATGATCACCACTCCGCGCTCCCAAAGTCT	4257
Db		4626	CAGGCTGCTCTTGAACCTCAAGATGATCACCACTCCGCGCTCCCAAAGTCT	4685
OY		4258	GGGATT-----ACAGTGTCCAGCCACATGACACAGCCACAGTGTACATTTTTTAAAT	4311
Db		4686	GGGATTACAGGACACAGGGGTAGCGCACTACACCTGGCTATCTCATCTTATATGTT	4745
OY		4312	TATTTT-----TTAATTAATGTTTATCTAAGG	4341
Db		4746	TCTTGAATAAGTATCTTAGAGTAGATTAACAGATCAGACAGACACAAGTTTGTATGGC	4805
OY		4342	CAGTAGACGTGACCTGCGT-----CTGTAAATCCAGCATTTAGAGGGCCCAAGTGC GG	4395
Db		4806	TGGGTGCGGTGGCTCAGCTGTGTGCTGCTTAATCCAGATCTTTAGAGAGGCCMAAGTGGC	4865
OY		4396	GGATCACTTGAACCTGGGAGTT-----CACGTGGGCAATATGTAGAACCCGTCCT	4449
Db		4866	AATTCATCTGAGCTCAGGAATTCGAGACCAAGCTGTGGCAATGGCAAAACCCATCTCT	4925
OY		4450	ACCAAAAATTTAAAAATTTAGCTGGAGTGTGGCATTTGCTGTGTTCCAGCTATTTG	4509
Db		4926	A--MAGAAATACAAAAATTTAGCCAGATGTGGTGGTGGCTGTAGTCTTAGCTACTTA	4983
OY		4510	GGAACTGAGTGTGGGGATGTGCTGAAGCTGTGAGTGCAGAGCTGAGTATGAT	4569
Db		4984	GGAGCTGGGGTGGGATCAATTCAGCCAGAGGATGGAAGTCACTGAGTGTGAT	5043
OY		4570	CACACGACGACTTGAAGCTG-----AGTGAAGGCTATCTCAAAAGCAAAACAAA	4621
Db		5044	TGCACATGGCACTCCAAGCTGGGCTCAAAATGAGATCTCTTCCAAAACAAAAGAA	5103
OY		4622	TAAATTTTATCTAAAGCT	4640
Db		5104	TACAAATCTTAAAGCT	5122
 RESULT 10 US-08-451-778A-7 Sequence 7, Application US/08451778A Patent No. 5830649 GENERAL INFORMATION: APPLICANT: Bergma, Derk J. APPLICANT: Stambolian, Dwight TITLE OF INVENTION: Human Galactokinase Gene NUMBER OF SEQUENCES: 33 CORRESPONDENCE ADDRESS: ADDRESSEE: Smithline Beecham Corp./Corporate ADDRESS: Intellectual Property STREET: 709 Swedeland Road/DW2220 CITY: King of Prussia STATE: Pennsylvania COUNTRY: USA				

ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,778A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Eagle, Alisha M.
REGISTRATION NUMBER: 37,126
REFERENCE/DOCKET NUMBER: P50268-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5364
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-451-778A-7

Query Match	6.08;	Score 301.4;	DB 2;	Length 7676;
Best Local Similarity	52.6%;	Pred. No. 9.2e-72;		
Matches 1052;	Conservative	0;	Mismatches 841;	Indels 106; Gaps 14;
Qy	2712	ATTTATTTATTTATTTATGACTTTTAAGACAGAGTTTTGCTCTTCTTACCCAGGCTTGATGCA	2771	
Db	3160	ATTTCGCCGTGTTTTTTTTTTTGGACACAGAGTTTCGGCTGTGTACACCATATGCTGAAATGCA	3219	
Qy	2772	ATGGCGGATCTCAGCTCACTGCAACCTCCACCTCCGGGATCAAGAGATTCTCGGCT	2831	
Db	3220	ATGGGCGCATCTCAGCTCACTGTAACCTCTGCTCCCGGGTTCAAGTGAATTCCTCGGCT	3279	
Qy	2832	CAGCTCTCTGAGTAGCTGAGTATACAGGCAACCCGACACGCGCTGATATTTTTTTGTA	2891	
Db	3280	CAGCTCTCCGAGTAGCTGAGTACAGAGGCGCCGCACTACGCTCGGCAATTTTTTATA	3339	
Qy	2892	TTTTTAGTAGAGAC-AGGGTTTCACCATGTTGGCCAGGCTGCTGCAGACCTT	2950	
Db	3340	GTTCTAGTAGAGACAGTGGGTTTTTCAACATGTTGGCCAGGCTGCTGCAGACCTT	3399	
Qy	2951	AGCTATATCACCTGCTCGACTGCTCCCAAGTGTGTGGATTATPAGCATGAGCCACTGCG	3010	
Db	3400	AGGTATCTCTCCGCTCAGCCTTCACAGTGTGTGGATTACAGGCGTAGACCATATGTGC	3459	
Qy	3011	C-----CAGTGAATTATAGAAAGTTAAAGGACATATGGCAATGCACACGCTATCTA	3060	
Db	3460	CGGTCTCTTTTTTTTTTTTTTTTAACTAAACATATCTCAGAAACCCAGAACCTATCTT	3519	
Qy	3061	CGTCTTCCTGCCAAGCAAGGAGAGGCTCGGGCTCACATTCTTGCGTTCCTACTTCC	3120	
Db	3520	ATCTTATGGCATGAAGGACATATCTTCGGCGTGCTCTTTTTTTTTTTTTTTTTTTTTT	3579	
Qy	3121	AAAAGCATTCAAGAACTGGCAGAGGCGCTTGAGACACCATTCATCCACTCTCAAGGTCCT	3180	
Db	3580	TGGGCGAGGTGGAGGCTTGCCCTGTTGCCACAGCTGTGAGTGCAGCGGCGCATCTCGGTT	3639	
Qy	3181	ATGGAGAGTTGAGGTCTCAGAGCAGGGAAGGCTCCGAGACAGGCTGTGACACAGGCGCTTG	3240	
Db	3640	CACGTGATCTTCCTCAGCTCTCGGCTCCAAATGATCTCTGCTTACGTTCCGATAGGT	3699	
Qy	3241	ATCCCTAACAACCCCAATCGGTGTCCTCTTACACAGGACCAACCCCACTGCTGCAG	3300	
Db	3700	GCGATTACTGGAACCCACACGAGCCCAATTTTTTATATTTTATAGTAGAGCGGG	3759	

QY 3301 CCCAGCTGGGCGATGACCATCATTACACAGCCAAAGTGGCTAATGCCCGCTTAGGCT 3360
 DB 3760 TTTTATAG-TTGGCCAGGCTGGGCTGCATCTCTGACCTGTGATCTGCCCGCTCAGCCT 3818
 QY 3361 CTTTCTCCCGCTGCTGCTGCTGGGCGGAGCATCTACAAGCTG---CTAATGGCG 3417
 DB 3819 CCCAATGTCTAGGATTTACATGTGTAGGACACTGACCTGGCTCCGTGTGGCTTTTAA 3878
 QY 3418 AGTCTCAATCTTCTGCTGCTGCTACTACATCATCCGCTTATTTATAGTAAGCTGGC 3477
 DB 3879 AGTCTCACATATTTTATAGCATTCAGGTGCTGTCTCAATTTACTTAATTTCTGTATACA 3938
 QY 3478 AGGGCTGGGGGGGGGGGGGGGGAAGATGTGGCTGGGGCTGGGAGCTGGGAGCTCTCG 3537
 DB 3939 CCTCACACTGCGATTTTACTTTCTTATTTATCTTTTATTTATTTATTTATTTATTTAT 3998
 QY 3538 GGGCTCTCCAGCCAGCTCAGGGCCAGTGCACAGTCCACTACACACATAGCTGGGCTC 3597
 DB 3999 TGAGACAGAGATCTTGGCTCTGTACCCAGGCTGGAGTGCAGATGCGCTCAGT 4058
 QY 3598 CTGACCAAGCTCTGGGCGACTGAGCTGAGGCTGGGGCTGGGGGCGAGAGTAAGA 3657
 DB 4059 GCAACCTGTCCCTCCAGGTTCAAGTATCTCTGCTCAGCCTCTG-----AGT 4110
 QY 3658 AGTACACACTGAGAGGCTGCTCAAGCCAGCCAGAGGGTTTATAGCCACCTTCTCCAC 3717
 DB 4111 AGCTAGATTTAGAGGATGTGGCCACACACACTGGCTATTTGATTTTATAGAGAT 4170
 QY 3718 CCCAGAGAGACCCCTGGAGCCAGGCTTGTCTGGCCCACTACTGCTGCTGTTTACT 3777
 DB 4171 GAGGTTTATACATTTGGTGGGCTGGCTGTGAACCTGCTGATCTGGCCACCT 4230
 QY 3778 GAATCCACACAGACTATAGGCGCCACATAGTACATTAATAAAGAGAGAGAGAGAG 3837
 DB 4231 CAGCTCCCAAGACTGGGATGACAGGATGAACCACTGGCTGGCCATCTTTTAT 4290
 QY 3838 AGAGAGAGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3897
 DB 4291 TTTTAAAGAGATGGTCTCTAAGTTGCCAGGCTGAGCTGATCTTGGGCTCAG 4350
 QY 3898 CAATCCCCCTGCTTACCTCCCAAGGGGCTGGGATTAAGAGTGTGAGCTACTGACATG 3957
 DB 4351 TAACTCTCTCAC--TAGTCTCTGGTGAAGTGAACCAAG----- 4391
 QY 3958 ACCAACAAGATGACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4017
 DB 4392 -----CACCGGTTTATCTGATCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4446
 QY 4018 ACCAGGCTGAGTGAAGTGGGGCAATCTTGGCTCACTGATACCTGTGCTCCAGGTG 4077
 DB 4447 GCCAGGCTGAGGCGAGT-GGCGGATCTGGGCTCACTGCAACCTCCGTCTTCAAGGTT 4505
 QY 4078 CAGGAGATTTCTGCTGCTTACCTCCCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4137
 DB 4506 CAAAGCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4565
 QY 4138 CTTGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4197
 DB 4566 GCGAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4625
 QY 4198 CAGGCTGCTTGAACCCCTGAGCTCAAGTATCAACCCAGCTGGCTCCCAAGTGT 4257
 DB 4626 CAGGCTGCTTGAACCCCTGAGCTCAAGTATCAACCCAGCTGGCTCCCAAGTGT 4685
 QY 4258 GGGATT-----ACAGGTGTCAGCAACCATGCAAGCCCAATGATTTTAAAT 4311
 DB 4686 GGGATTACAGGACAGGCTGAGCACTACCTGCTGATCTGCTGCTGCTGCTGCTGCTG 4745
 QY 4312 TATTTT-----TATTAATTAATGTTTATCTAAGGC 4341
 DB 4746 TCTTAGAATGATTTCTTAGAGTAGATTTACAGAGTCAAGAGACACAGATTTTGAAGC 4805
 QY 4342 CAGTAGCACTGACTGCGGT-----CTGTAATCCAGCACTTTGAGGGCCAAAGTGGC 4395

DB 4806 TGGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4865
 QY 4396 GATATCATTTGAGCTGAGGATTT-----CAGCGTGGCAACATAGTAGAGACCCGCTCT 4449
 DB 4866 AGATTCATTTAGCTAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGG 4925
 QY 4450 ACCAAAAATTTAAAAATTTAGCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4509
 DB 4926 A--AAGAAATACAAAAATTTAGCCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4983
 QY 4510 GGAAGCTGAGTGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4569
 DB 4984 GAGGCTGGGGTGGGAGGATCAATTGAGCCAGAGGTTGAGACTGAGCTGAGCTGAT 5043
 QY 4570 CACACACTGACACTTCAAGCTG-----AGTACAGGCTATCTCAAAAGCAACAAA 4621
 DB 5044 TGCACCATGCACTGCAAGCTGGGCTCAAGTGAAGATCTGCTGCTGCTGCTGCTGCT 5103
 QY 4622 TATGTTTATCTAAGCT 4640
 DB 5104 TACAGTATCTTAAAGCT 5122

RESULT 11
 US-08-998-208-7
 : Sequence 7, Application US/08998208
 : Patent No. 5880105
 : GENERAL INFORMATION:
 : APPLICANT: Bergsma, Derek J.
 : APPLICANT: Stambolian, Dwight
 : TITLE OF INVENTION: Human Galactokinase Gene
 : NUMBER OF SEQUENCES: 33
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: SmithKline Beecham Corp./Corporate
 : ADDRESSEE: Intellectual Property
 : STREET: 709 Swedeland Road/UM2220
 : CITY: King of Prussia
 : STATE: Pennsylvania
 : COUNTRY: USA
 : ZIP: 19406-0939
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/998,208
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/451,777
 : FILING DATE: 26-MAY-1995
 : APPLICATION NUMBER: PCT/US94/10825
 : FILING DATE: 23-SEP-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Eagle, Ailsa M.
 : REGISTRATION NUMBER: 37,126
 : REFERENCE/DOCKET NUMBER: P50268-1B
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 610-270-5364
 : TELEFAX: 610-270-5090
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 7676 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : US-08-998-208-7

Query Match 6.0%; Score 301.4; DB 2: Length 7676;
 Best Local Similarity 52.6%; Fred. No. 9.2e-72;

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESS: Intellectual Property
STREET: 709 Swedeland Road/DW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sutton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P50268-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5024
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-0595-06743-7
Query Match 6.0%; Score 301.4; DB 5; Length 7676;
Best Local Similarity 52.6%; Pred. No. 9.2e-72;
Matches 1052; Conservative 0; Mismatches 841; Indels 106; Gaps 14;

3181 ATGGAGAGTTGAGTCCAGACGAGGAAGGTCCTGACAGGCTGACAGGAGGCTCTG 3240
3640 CACTGATCTCCACCTCTGCTGCTCCAAATGATCTCTGCTTACCTCTGATAGT 3699
3241 ATCCCTACAAACCCCAATGCGTCCCTCTACAGAGACCCAAAGCCACTGCTCAG 3300
3700 GGGATTACTGGAACCCACACACAGCCCGACCAATTTTATTTTATGATGAGACGGG 3759
3301 CCCACTGCTGGCCATGACATCACTTACAAAGCCAAAGTGGCTAATGCCGCTAGCT 3360
3760 TTTCATG-TTGGCAGAGCTGGCTCGAATCTGATCTGATCTGATCTGATCTGATCT 3818
3361 CTTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3417
3819 CCATATGCTAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3878
3418 AGTTCCTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3477
3879 AGCTCCAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3938
3478 AGGGCTGGCGGGGGGCTGGGAAGATGCTGGGGCTGGGAGCTGGAGCTCCTGG 3537
3939 CCTCACACTGCAATTAATCTTCTTATTTATTTATTTATTTATTTATTTATTTATTT 3998
3538 GGGCTCCAGCCCAAGCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3597
3999 TGAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4058
3598 CTGACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3657
4059 GCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4110
3658 AGTCACAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3717
4111 AGTAGATTAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4170
3718 CCCAGAGGAGCCCTGGAGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3777
4171 GAGGTTTATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4230
3778 GAATCCCAACAGACTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3837
4231 CAGCTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4290
3838 AGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4350
4291 TTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4391
3898 CAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3957
4351 TAACTTCTCACG-TAGTCTCTGGTATGATGATGATGATGATGATGATGATGATGATGAT 4391
3958 ACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4407
4392 -----CACCGGTTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4446
4018 ACCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4077
4447 GCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4505
4078 CAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4137
4506 CAGCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4565
4138 CTTGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4197
4566 GCGAGTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4625
4198 CAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4257
4626 CAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4685
4258 GGGATT-----ACAGTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4311

```

Db 4686 GGGATTACAGGCACAGCGCTGAGGCACTACACCTGGCGCTATCTCATTCTCTTAAGT 4745
QY 4312 TATTT-----TTATTAATGTTTATCTAAGC 4341
Db 4746 TCTTAGAAATGATCTTAGAGATGAGATTACAGAGTCAAGAGACACAAGTTTGTAGC 4805
QY 4342 CAGTAGAGTACTCGGGT-----CTGTATCCAGCACTTTAGAGGCCAAGTCCG 4395
Db 4806 TGGGTGGGTGCTACAGCTCTGTCTGTATTCCTAGTACTTTAGAGGCCAAGTGGC 4855
QY 4396 GGATCACTTGAGCCCTGGAGTT-----CAGCGTGGGCACATAGTAGAGACCCGCTCT 4449
Db 4866 AGATTTCATTAGTCTAGGAATTCAGAGACCAAGCTGGGCAACATGGCAAAACCCATCTCT 4925
QY 4450 ACCAAAAATTTAAAAATTTACTGGAGTGGTGGCATTTGGCTGTGGGCCAGTACTTG 4509
Db 4926 A--AAGAAATCAAAAATTTACCCAGGTGGTGGTGTGTGCTGTAGTCTTAGCTACTTA 4983
QY 4510 GGAAGCTAGTGTGGGGATGGCTGTGAAGCTGTGAGGTGAGAGTGCAGTANGAT 4569
Db 4984 GGAGCTGGGTGGAGAGATCAATTTAGCCAGAGAGGTGAGACTGCAGTGCTGAT 5043
QY 4570 CACACCACGACTTCACCTG-----AGTGCAGGCTATCTCCAAACCAACAAA 4621
Db 5044 TGCACCAAGCAGCATCTCCAGCTGGGCTCAAGTGTGATCTGTCCAAAAACAAAAAGA 5103
QY 4622 TAATGTTATCTTAACGGT 4640
Db 5104 TACAGTATCCTTAAGCT 5122

RESULT 13
US-08-724-394A-20/c
Sequence 20, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Therefo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Flets, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs

```

	Type:	nucleic acid
	STRANDEDNESS:	not relevant
	TOPOLOGY:	not relevant
	MOLECULE TYPE:	cDNA
	FEATURE:	
	NAME/KEY:	misc_feature
	LOCATION:	1..246240
	OTHER INFORMATION:	/note= "HLA-H.CONTIG"
US-08-724-394A-20		

Query Match	Best Local Similarity	Score	Pred.	No. Matches	Mismatches	Indels	Gaps
3974	TTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGCTTCATCCTCATCACCCAGGCTGGAGTGC	6.0%; 299.8;	72.2%; 0;	DB 2; Length 246240;			
3992	TTTTTTTTTTTTTTTTTTTTTTTGGAGACGGAAGTCTGTCTGCGGCCAGGCTGGAGTGC						
4034	AGTGGGGGCAATCTGGGCTCAGCTGTAACTCTTGCCCTCCAGGTGCAAGGATTCTCTGCG						
3332	AGTGGGCGG-GATCTCGGCTCAGTGAAGCTCGGCTCCGGGTTACAGGCATTCTCTGCG						
4094	CTTAGCCTCTCGATGACTGGAATTTATAGGACACACACACACAGCCTGGCTAAATTTTTT						
3173	CTCAGCCTCCCAAGTAAGTCTGGAGCTACAGGCGGCCGACACTAGCCCGGCTAAATTTTT--						
4154	TTTTTCTGTATTTTATAGTGAAGACAGGCTTTATCATGTGTGGCCAGGCTGGCTGAAGC						
3115	-----TGTATTTTTTGTAGAGACGGGGTTTACCAGTTTACCGGGAGTGTCTGATTC						
4214	CCCGTACCTCAAGGATTCACACCCACACCTCGGCTCCCAAAGTCTGGGATTACAGTGTCA						
3062	TCTGACCTC-GTATCCGCCCCCGCTCGGCTCCCAAAGTGTGGGATTACAGGCGTGA						
4274	GCCACCATGACACAGCCACATGTGTCAATTTTTAAAA-----TTATTTT						
3004	GCCACCGGCGCCGCTTATPAAAAGACTTTTATAGTAAAAACAAGATCTATATCTTAATG						
4318	TTAATTAAAAATGTTATATCTAAGGCAAGTACAGTACGTCGCTGTATATCCAGCACTT						
2944	TATGATTATTAAGAATAAACCCCGGCTGGCGCAGTGGCTCACACTGTATATCCAGCACTT						
4378	TGAGGGGCAAGTGTGGGGATCACTTGAGGCTGGGAGTT-----CAGCGTGGGCACA						
2884	TGGGAGGCGGAGGGGGGTATGATCACTTGAGGTGAGAGTTGAGACCAAGCCTGGGCCAACA						
4432	TAGTGAGACCCGCTCTCTACCAAAAAATTTAAAAAATTAAGTGGGAGTGGGCACTTGGC						
2834	CGGCCAAACCTGTCTCTACTTAAAAA-TACAAAAAATTAAGTGGGCAAGTGGGAGGCACC						
4432	TGTGTGCCCACTACTTGGGAAGCTGAGAGTGTGGGGATGGCTGAAGCCTGTGAGGTCGAG						
2765	TGTATATCCCACTGTTGGGAGACTGAGGACGAGGAATCGTTAACTGGGAGTGGGG						
4552	GCTGCATGAGCTATGATCACACCACTGCACTTAAAGCTGAGTATACAGCGCATCTCAAA						
2705	ATTGCAGTGAAGCAAGACTGCAACCACTGCACTGAGCTGGGCAAGAGGCTGAGACAAAG						
4612	GCAAAACAAATPA 4624						
2645	TCTCAAAATPA 2633						

RESULT 14

US-08-724-394A-21/c

Sequence 21, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Krommal, Gregory S.

APPLICANT: Laufer, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchinashi, Zenta
 APPLICANT: Wolff, Roger K.
 TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
 TITLE OF INVENTION: Sequences and Antibodies thereto
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/724,394A
 FILING DATE: 01-OCT-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Flits, Renee A.
 REGISTRATION NUMBER: 35,136
 REFERENCE/DOCKET NUMBER: 017957-000100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 246240 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..246240
 OTHER INFORMATION: /note= "HLA-H CONTIG"
 JS-08-724-394A-21

Query Match	6.0%	Score 299.8	DB 2	Length 246240
Best Local Similarity	72.2%	Pred. No. 2.3e-70		
Matches 486	Conservative	0	Mismatches 152	Indels 35
			Gaps	6
Qy 3974	TTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGGTTTCACCTCATCACCAGGCTGAGATGC	4033		
Db 3292	TTTTTTTTTTTTTTTTTTTTTTTTTATAGACGAGATCTTGCTGTGTGCCAGGCTGAGATGC	3233		
Qy 4b34	AGTGGGGCAATCTTGCTCAGTAACTTGCCTCCACAGTGCAGACGATTTCTCTGC	4093		
Db 3232	AGTGGCGG-GATCTCGGCTCACTCAAGATCCGCTCCGGGTTCAAGCCATTTCTCTGC	3174		
Qy 4094	CTTAGCCCTCCGAGTAGCTGGAATTAAGGACACACACAGCCCTGGGCAATTTTTTTT	4153		
Db 3173	CTCAGCCCTCCAGATAGCTGGACTACAGGGCGCCGCGCATACGCCGCGCTAATTTTT--	3116		
Qy 4154	TTTTTCTGTATTTTATAGTAGAGACAGGGTTTCATCATGTGGCCAGGCTGGTCTGAAC	4213		
Db 3115	-----TGTATTTTNTATAGACAGGGGTTTCAACGGTTTAAAGCCGGGATGCTCTGATC	3063		
Qy 4214	CCCTGACCTCAAGTATCATCACCCACCTCGGCTCCCAAGTCGTGGGATTTACAAGTGTCA	4273		
Db 3062	TCTGTACCTC--GTGATCCGCGCGCTCGGCTCCCAAGTGTGGGATTTACAAGCGTGA	3005		
Qy 4274	GCCACATGACACAGCCCAATGATGTATTTTTTAAA-----TTATTTT	4317		
Db 3004	GCCACCGCGCCGCGCTTATTAAGACTTTTATGATATAAACAAGATCTATATCTCTAATG	2945		
Qy 4318	TTAATTAATAATGTTTATCTAAGGCCAGTAGCAGTCTCGGCTGTATATCCACGACTT	4377		
Db 2885	TATGATTTAAGATATACCCGGGCTGGGCGAGTGGCTTCAACACCTGTATCTCCAGCACTT			

	4378	TGAGGGGCGAAGTGGCGGGATCACTTGAGCCTGGAGTT- - - - -	CAGCGTGGGCAACA	4431
QY	4378	TGAGGGGCGAAGTGGCGGGATCACTTGAGCCTGGAGTT- - - - - <td>CAGCGTGGGCAACA</td> <td>4431</td>	CAGCGTGGGCAACA	4431
Db	2884	TGGGAGGCGCGAGGGGTAGATCACTTGAGGTCAGGAGTTGAGACACAGCCTGGCCAAACA	2825	
QY	4432	TAGGAGAGCCCGCTGTCTACCAAAAATTAAATAATAGCTGGGAGTGGTGGCAATTTGCC	4421	
Db	2884	CGCGGAAACCTGTCTCTACTATAAAAA-TACAAAAATTATAGTGGGCAATGGTGGCAGGCACC	2766	
QY	4492	TGTGGTCCCACTACTTGGGAAGCTGAAGTGTGGGGATGGCTGAACCTGTGAGAGTGCAG	4551	
Db	2765	TGTATATCCCACTCTTGGGAGACTGTAGGCGCAGGAAATCGTTGAACCTGGGACGTGGGG	2706	
QY	4552	GCTGAGTGAAGCTATGATCACACCACTGCACCTTGAAGTCAGCAGGCTATCTCAAAA	4611	
Db	2705	ATTGCAAGTGAACCAAGACTGCACACCTGCACTCCAGCCTGGGCGACAGGGTGAGACAACG	2648	
QY	4612	GCAAAACAAATTA 4624		
Db	2645	TCTCAAAAAATTA 2633		

```

      RESULT 15
US-08-724-394A-22/c
Sequence 22, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchinashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Flits, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.COMTIG"
US-08-724-394A-22

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:29:22 ; Search time 1005.38 Seconds
(without alignments)
11199.691 Million cell updates/sec

Title: US-09-622-964-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaataattgtctctctg.....tttgaggagcgagcgaggt 5000

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4990	99.8	16125	AA221226	Human CGICE genom
2	2758.2	55.2	16550	ABAI4559	Human nervous syst
3	2758.2	55.2	18530	ABAI4557	Human nervous syst
4	2758.2	55.2	18537	ABAI4558	Human nervous syst
5	1291.8	25.8	6396	AB134247	Human immune syste
6	1284.2	25.7	6396	AB134246	Human immune syste
7	375.6	7.5	17792	AA532727	Human genomic DNA
8	375.6	7.5	17792	AA536099	Human cardiovascular
9	356.2	7.1	2816	AAK83109	Human immune/haema

C 10	352.6	7.1	15413	22	AAK84002	Human immune/haema
C 11	351.4	7.0	201143	24	ABK83568	Human DNA differen
C 12	351.2	7.0	32224	22	AAK89986	Human digestive sy
C 13	350.6	7.0	13819	22	AA136231	Human muscleloxele
C 14	349.6	7.0	30393	22	AAK67239	Human immune/haema
C 15	344.4	6.9	1979	22	AAH17407	Human immune/haema
C 16	340.6	6.8	13821	22	AA136230	Human muscleloxele
C 17	338.6	6.8	9397	22	AAK78065	Human immune/haema
C 18	337.6	6.8	149480	24	ABK61947	Human immune/haema
C 19	337.6	6.8	149480	24	ABK61948	Human immune/haema
C 20	337.6	6.8	149480	24	ABK61949	Human immune/haema
C 21	336.6	6.7	10139	22	ABK68365	Human immune/haema
C 22	335.2	6.7	19632	22	AAK529226	Human immune/haema
C 23	335	6.7	6451	22	AAK72461	Human immune/haema
C 24	334.6	6.7	22481	17	AAH11658	Human immune/haema
C 25	334.6	6.7	22484	24	ABK72461	Human immune/haema
C 26	334.6	6.7	22484	24	ABK72461	Human immune/haema
C 27	334.6	6.7	22484	24	ABK72461	Human immune/haema
C 28	334.6	6.7	22484	24	ABK72461	Human immune/haema
C 29	332.8	6.7	23825	22	AAK79161	Human immune/haema
C 30	332.2	6.6	139904	24	ABK83562	Human immune/haema
C 31	331.6	6.6	32194	22	AAK71591	Human immune/haema
C 32	330.2	6.6	11301	22	AAK71591	Human immune/haema
C 33	330	6.6	939	22	AAK83508	Human immune/haema
C 34	329	6.6	10139	22	AAK529226	Human immune/haema
C 35	329	6.6	175737	24	ABK83571	Human immune/haema
C 36	328.4	6.6	220895	24	ABK84798	Human immune/haema
C 37	328	6.6	14448	22	AAK71815	Human immune/haema
C 38	328	6.6	14448	22	AAK71811	Human immune/haema
C 39	328	6.6	14448	22	AAK87580	Human immune/haema
C 40	328	6.6	14448	22	AAK62922	Human immune/haema
C 41	328	6.6	14451	22	AAK71818	Human immune/haema
C 42	328	6.6	14451	22	AAK73116	Human immune/haema
C 43	328	6.6	14451	22	AAK87587	Human immune/haema
C 44	328	6.6	14451	22	AAK62925	Human immune/haema
C 45	326.4	6.5	14426	22	AAK71817	Human immune/haema

ALIGNMENTS

RESULT 1	AA221226	standard; DNA; 16125 BP.
ID	AA221226	
XX	AA221226;	
AC	22-NOV-1999	(first entry)
XX		
DT	22-NOV-1999	(first entry)
XX		
DE	Human CGICE genomic DNA sequence.	
XX		
KW	CGICE; Best's macular dystrophy; mutation; diagnosis; detection;	
KW	BMD; age-related macular dystrophy; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Intron	1..1513
FT		/*tag- a
FT	exon	/*number- 1
FT		1514..1570
FT		/*tag- b
FT		/*number- 1
FT	Intron	1571..3278
FT		/*tag- c
FT	exon	/*number- 2
FT		3279..3466
FT		/*tag- d
FT		/*number- 2
FT	CDS	3315..15787
FT		/*tag- e
FT		/*product- "CGICE protein"
FT		/*note- "contains introns"

FT	Intron	3467..6616
FT		/*tag= f
FT		/number= 3
FT	exon	6617..6711
FT		/*tag= g
FT		/number= 3
FT	Intron	6712..7227
FT		/*tag= h
FT		/number= 4
FT	exon	7228..7461
FT		/*tag= i
FT		/number= 4
FT	Intron	7462..8353
FT		/*tag= j
FT		/number= 5
FT	exon	8354..8508
FT		/*tag= k
FT		/number= 5
FT	Intron	8509..8896
FT		/*tag= l
FT		/number= 6
FT	exon	8897..8974
FT		/*tag= m
FT		/number= 6
FT	Intron	8975..9676
FT		/*tag= n
FT		/number= 7
FT	exon	9677..10032
FT		/*tag= o
FT		/number= 7
FT	Intron	10033..11212
FT		/*tag= p
FT		/number= 8
FT	exon	11213..11293
FT		/*tag= q
FT		/number= 8
FT	Intron	11294..11556
FT		/*tag= r
FT		/number= 9
FT	exon	11557..11708
FT		/*tag= s
FT		/number= 9
FT	Intron	11709..13918
FT		/*tag= t
FT		/number= 10
FT	exon	13919..14557
FT		/*tag= u
FT		/number= 10
FT	Intron	14558..15768
FT		/*tag= v
FT		/number= 11
FT	exon	15769..16125
FT		/*tag= w
FT		/number= 11
PN	WO9943695-A1.	
XX		
PD	02-SEP-1999.	
XX		
XX	22-FEB-1999;	99WO-US03790.
XX		
PR	25-FEB-1998;	98US-0075941.
PR	18-DEC-1998;	98US-0112926.
XX		
PA	(MERI) MERCK & CO INC.	
PA	(UYUP-) UNITV UPSPALA.	
XX		
PI	Petrukhin K, Caskey CT, Metzker M, Madellius C;	
XX		
DR	WPI. 1999-540560/45.	
DR	P-PSDB; AA129953.	
PT	Human and mouse polynucleotides encoding CGICE polypeptides	

[illegible]

Db 781 CACCCCTCCACACAGCCCGGAGGCTTCTCTTCATCTCTGAAACAATTACTGA 840
OY 841 TGGGCGCTCCGACCAATACACAGCCAGATTAACGTAATGTCACACAGCCCAATCAGA 900
Db 841 TGGGCGCTCCGACCAATACACAGCCAGATTAACGTAATGTCACACAGCCCAATCAGA 900
OY 901 GCTCCCGTCAGCATATGCAAGATTCTGTCATTTTCTAGGGTGAATGAAATTCACAGCA 960
Db 901 GCTCCCGTCAGCATATGCAAGATTCTGTCATTTTCTAGGGTGAATGAAATTCACAGCA 960
OY 961 ACACATCTCTTTCAGATTAAGGCGACTGAGGCTGAGAGAGAGCTGAACCTAACCCGGGG 1020
Db 961 ACACATCTCTTTCAGATTAAGGCGACTGAGGCTGAGAGAGAGCTGAACCTAACCCGGGG 1020
OY 1021 TCACACACACAGGTGGCAAGGCTGGACCAAGAAACAGAGACTGTGACTGCAGCCGGT 1080
Db 1021 TCACACACACAGGTGGCAAGGCTGGACCAAGAAACAGAGACTGTGACTGCAGCCGGT 1080
OY 1081 ATTCAATCTTTCATAGCCACAGGCTGTCAAAAGACCCAGGGCCTAGTCAGAGGCTCC 1140
Db 1081 ATTCAATCTTTCATAGCCACAGGCTGTCAAAAGACCCAGGGCCTAGTCAGAGGCTCC 1140
OY 1141 TCTCTCTGAGAGATCTGTCGACAGAGATTGAAGCTCAGACAGCCCTTAACCCGCCA 1200
Db 1141 TCTCTCTGAGAGATCTGTCGACAGAGATTGAAGCTCAGACAGCCCTTAACCCGCCA 1200
OY 1201 CTCTCTCTGCAAGGCTCAGAGGGGTCAGACACAGTGTGAGCAGATCCTTTAGCCTGGA 1260
Db 1201 CTCTCTCTGCAAGGCTCAGAGGGGTCAGACACAGTGTGAGCAGATCCTTTAGCCTGGA 1260
OY 1261 TTTTAGGCGCATGTGTAGAGGGGGGTGTTCCCTTAATTCAGCCCTGTCTCAGCCCA 1320
Db 1261 TTTTAGGCGCATGTGTAGAGGGGGGTGTTCCCTTAATTCAGCCCTGTCTCAGCCCA 1320
OY 1321 CCCTCCAGAGAAATTAAGAGGGGCCAATGGCCAGGCTGTGCTAGCCGTCTTGAGCA 1380
Db 1321 CCCTCCAGAGAAATTAAGAGGGGCCAATGGCCAGGCTGTGCTAGCCGTCTTGAGCA 1380
OY 1381 GATTACAGAGAGGACTTAAGACAGAGACTCTTTGTGAGGTCCTGTGCTTAGGAGTCAA 1440
Db 1381 GATTACAGAGAGGACTTAAGACAGAGACTCTTTGTGAGGTCCTGTGCTTAGGAGTCAA 1440
OY 1441 GTGAGCGGCTCAGACTCAGCTGTGAGGAGTGCACAGCTCTTAAGAGTGGAGGGCACT 1500
Db 1441 GTGAGCGGCTCAGACTCAGCTGTGAGGAGTGCACAGCTCTTAAGAGTGGAGGGCACT 1500
OY 1501 GGCACAGAGTCCAGGAGGTCCACAGCTTAAGTGCACAGACTTCTGTGGATCATCG 1560
Db 1501 GGCACAGAGTCCAGGAGGTCCACAGCTTAAGTGCACAGACTTCTGTGGATCATCG 1560
OY 1561 GACCACCTGGAAACCCCACTGTGAGTACAAAGGTCCCCAGGTGAGTGGGGCTT 1620
Db 1561 GACCACCTGGAAACCCCACTGTGAGTACAAAGGTCCCCAGGTGAGTGGGGCTT 1620
OY 1621 TGAGGCTTCAAGGTTGATGAGGCTCTTGGCTATTTGTGTGGATTAAGCAGACAGAGC 1680
Db 1621 TGAGGCTTCAAGGTTGATGAGGCTCTTGGCTATTTGTGTGGATTAAGCAGACAGAGC 1680
OY 1681 AACACATGCGAGGTTGTGGGACCTGTGTCTGTGCAAAATGCCCTGAGGTGGAAATG 1740
Db 1681 AACACATGCGAGGTTGTGGGACCTGTGTCTGTGCAAAATGCCCTGAGGTGGAAATG 1740
OY 1741 ACCTTGCTGATCAGAGCAGACAGCCAGAGTGTGCTCAGCAAAACACAGGGA 1800
Db 1741 ACCTTGCTGATCAGAGCAGACAGCCAGAGTGTGCTCAGCAAAACACAGGGA 1800
OY 1801 AAGATGAGAGGGGCTCAATCACTACATCAAAATTAATTAAGAGTCCCTTAAGAAAAA 1860
Db 1801 AAGATGAGAGGGGCTCAATCACTACATCAAAATTAATTAAGAGTCCCTTAAGAAAAA 1860
OY 1861 GAAGGTCTCTCTTGTGATAGAGAGGAGAGAGGGGTTGTCTTAATTAATTAAG 1920
Db 1861 GAAGGTCTCTCTTGTGATAGAGAGGAGAGAGGGGTTGTCTTAATTAATTAAG 1920

OY 1921 GAGAGCCGCCCTCCAAATAAGGAGGAGAGACCCAGACCCCGGGGTGTGTGT 1980
Db 1921 GAGAGCCGCCCTCCAAATAAGGAGGAGAGACCCAGACCCCGGGGTGTGTGT 1980
OY 1981 TTCCAGGGGGAGCTCGAACCTTTAGAGGAGCGGTGGAGAACCGCTGTATTCAGGCTC 2040
Db 1981 TTCCAGGGGGAGCTCGAACCTTTAGAGGAGCGGTGGAGAACCGCTGTATTCAGGCTC 2040
OY 2041 TCGAGAGAAAGAGAGCGGCCGCCAAATAATCTCTCCCGGCGATGAAGAAATGTTGGC 2100
Db 2041 TCGAGAGAAAGAGAGCGGCCGCCAAATAATCTCTCCCGGCGATGAAGAAATGTTGGC 2100
OY 2101 CTCTCTCAAAAAGATGAAGAGAGAACCGAGTGTATGTGATATTTTAAACTCCA 2160
Db 2101 CTCTCTCAAAAAGATGAAGAGAGAACCGAGTGTATGTGATATTTTAAACTCCA 2160
OY 2161 GGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTTAGCGCTTCTAGAGAACCAAGA 2220
Db 2161 GGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTTAGCGCTTCTAGAGAACCAAGA 2220
OY 2221 GGAGCTTCATCTGAGAGAGAAACAGGAGAAACAGGACAGATATCTGTATTAATTCA 2280
Db 2221 GGAGCTTCATCTGAGAGAGAAACAGGAGAAACAGGACAGATATCTGTATTAATTCA 2280
OY 2281 AGTAGATTAAGTCTCTAGAAATATCAAGCAAGGTGAGAGACACAGACCCGGTG 2340
Db 2281 AGTAGATTAAGTCTCTAGAAATATCAAGCAAGGTGAGAGACACAGACCCGGTG 2340
OY 2341 GCAGTGGGCTCATTTCCAGGTTGGATGTGGAAACATCTTTTAAAGGAACTGG 2400
Db 2341 GCAGTGGGCTCATTTCCAGGTTGGATGTGGAAACATCTTTTAAAGGAACTGG 2400
OY 2401 AGTGGAGAAAGAACCAAGTATCTCAGAGAGCTTCTCAGAGAGAGATCAGCA 2460
Db 2401 AGTGGAGAAAGAACCAAGTATCTCAGAGAGCTTCTCAGAGAGAGATCAGCA 2460
OY 2461 GGTGAGAAAGCCCTGAGCCACCATTCAGTAAACATCATTTGACATCTTACCAGCTAG 2520
Db 2461 GGTGAGAAAGCCCTGAGCCACCATTCAGTAAACATCATTTGACATCTTACCAGCTAG 2520
OY 2521 GTTCATTAATGGAATATGGAATATGTTGAGAGAGGAGGCTGCTGCTTCCATCTT 2580
Db 2521 GTTCATTAATGGAATATGGAATATGTTGAGAGAGGAGGCTGCTGCTTCCATCTT 2580
OY 2581 CTGACACTAGGGTGTGAGAGAGCTTGGAGGCTAACGAACAAGATGGCTGAGAACACT 2640
Db 2581 CTGACACTAGGGTGTGAGAGAGCTTGGAGGCTAACGAACAAGATGGCTGAGAACACT 2640
OY 2641 GCCTAGCCCAAGAGACCTGAGCTTAATGTGTAGACATTTGCTGCTTACTGCTTGTG 2700
Db 2641 GCCTAGCCCAAGAGACCTGAGCTTAATGTGTAGACATTTGCTGCTTACTGCTTGTG 2700
OY 2701 TTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
Db 2701 TTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
OY 2761 GCTTGAAGTCAATGAGGATCTGAGTCACTGCAACCTTCACCTCTCGGATTAAGGA 2820
Db 2761 GCTTGAAGTCAATGAGGATCTGAGTCACTGCAACCTTCACCTCTCGGATTAAGGA 2820
OY 2821 TTCTCTGCTCAGCTCTGAGTATGCTGGGATTAACAGGCAACCCGACACAGCTGATTA 2880
Db 2821 TTCTCTGCTCAGCTCTGAGTATGCTGGGATTAACAGGCAACCCGACACAGCTGATTA 2880
OY 2881 ATTTTGTGATTTTATTAAGAGAGGTTTACACATGTTGGCAGGCTGTCTGAAAC 2940
Db 2881 ATTTTGTGATTTTATTAAGAGAGGTTTACACATGTTGGCAGGCTGTCTGAAAC 2940
OY 2941 TCTGACCTTAAGTATCAGCTGCTGACTTCCCAAGTGTGGATTAAGGCAATGA 3000
Db 2941 TCTGACCTTAAGTATCAGCTGCTGACTTCCCAAGTGTGGATTAAGGCAATGA 3000

QY	3001	GCACATGGCCCCATGATATTATAGAAAGTTAAAGGCACATGGCCATATGCACACGCTATATCTA	3060
Dp	3001	GCACATGGCCCCATGATATTATAGAAAGTTAAAGGCACATGGCCATATGCACACGCTATATCTA	3060
QY	3061	CGTCTCCCTGCGCAAAAGCAAAAGGGGAGCGCTGGGGCTCACTTCTGTGGGTTCTCACTTCC	3120
Dp	3061	CGTCTCCCTGCGCAAAAGCAAAAGGGGAGCGCTGGGGCTCACTTCTGTGGGTTCTCACTTCC	3120
QY	3121	AAAAGCAGTGAACATGGCAGAGGGCGTTGGAGACACTTTCATCCACGCTCTTAGGGTCCCT	3180
Dp	3121	AAAAGCAGTGAACATGGCAGAGGGCGTTGGAGACACTTTCATCCACGCTCTTAGGGTCCCT	3180
QY	3181	ATGGGAGACTTGAAGTCCAGACAGAGGAGGGTCTCTGACAGGCTCTGACCAGGGCCCTCTG	3240
Dp	3181	ATGGGAGACTTGAAGTCCAGACAGAGGAGGGTCTCTGACAGGCTCTGACCAGGGCCCTCTG	3240
QY	3241	ATCCCTACAAACCCCAATGGGTGTCCTCTACACAGGACCCAAAGCCCAACCTGCTCTCAG	3300
Dp	3241	ATCCCTACAAACCCCAATGGGTGTCCTCTACACAGGACCCAAAGCCCAACCTGCTCTCAG	3300
QY	3301	CCCACTGGCTGGCATGACATCACTTACACAACGCAAGTGGTCAATGCCCCGTTAAGCT	3360
Dp	3301	CCCACTGGCTGGCATGACATCACTTACACAACGCAAGTGGTCAATGCCCCGTTAAGCT	3360
QY	3361	CCCTTCCTCCGCTGCTGCTGTGTGGGGGGGAGCATCTACAACTGCTATATGGGAGT	3420
Dp	3361	CCCTTCCTCCGCTGCTGCTGTGTGGGGGGGAGCATCTACAACTGCTATATGGGAGT	3420
QY	3421	TCCATATCTCTGCTGCTGCTACATACATCAATCCGCTTATTTATATAGTAAAGCTGGCAGG	3480
Dp	3421	TCCATATCTCTGCTGCTGCTACATACATCAATCCGCTTATTTATATAGTAAAGCTGGCAGG	3480
QY	3481	GCTGGGGCGGGGGGCTGGGAAGATGTCGTGGGGGCTGGGAGCTGGGAGCTCCTGGGGG	3540
Dp	3481	GCTGGGGCGGGGGGCTGGGAAGATGTCGTGGGGGCTGGGAGCTGGGAGCTCCTGGGGG	3540
QY	3541	CCTCCACAGCAGCTCAGGGGCCAGTGCACAGTCCACTACACAACTAAAGCTGGGCTCTG	3600
Dp	3541	CCTCCACAGCAGCTCAGGGGCCAGTGCACAGTCCACTACACAACTAAAGCTGGGCTCTG	3600
QY	3601	ACCAAGCTCCTGGGCACTGGAGCTGAGGGCTGGGGCGCTGGGGGCTGGGCGAGATTAAGAAAT	3660
Dp	3601	ACCAAGCTCCTGGGCACTGGAGCTGAGGGCTGGGGCGCTGGGGGCTGGGCGAGATTAAGAAAT	3660
QY	3661	CACACTGAGAGGCTGCTCAAGCCAGAGGCGAGAGGGTTTATCCACCCCTCTCCAAACCC	3720
Dp	3661	CACACTGAGAGGCTGCTCAAGCCAGAGGCGAGAGGGTTTATCCACCCCTCTCTCCAAACCC	3720
QY	3721	AGGAGGACCCCTGGAGCCCAAGGCTTGTGTGTCGGCCCACTCACTGAGGCGCTGTTTACTGAA	3780
Dp	3721	AGGAGGACCCCTGGAGCCCAAGGCTTGTGTGTCGGCCCACTCACTGAGGCGCTGTTTACTGAA	3780
QY	3781	TCCACACAGAGCTCATAGGCCCACTAATATATTAATAAAGAGAGAGAGAGAGAGA	3840
Dp	3781	TCCACACAGAGCTCATAGGCCCACTAATATATTAATAAAGAGAGAGAGAGAGAGA	3840
QY	3841	GAGAGAGATGAGTCTCACTGTGTGTCCAGAGCTGATCGCAACCTCAAGGCTCAAGCAA	3900
Dp	3841	GAGAGAGATGAGTCTCACTGTGTGTCCAGAGCTGATCGCAACCTCAAGGCTCAAGCAA	3900
QY	3901	TCCCCCTGCTTAAAGCTCCCAAGGGGGCTGGAGTTACAGGTGTGAGCTACTGCACATTGACC	3960
Dp	3901	TCCCCCTGCTTAAAGCTCCCAAGGGGGCTGGAGTTACAGGTGTGAGCTACTGCACATTGACC	3960
QY	3961	AACCAATGAGTACTTTTTTTTTTTTTTTTTTTTTTGTAGACAGGGTTTCACTCCATACAC	4020
Dp	3961	AACCAATGAGTACTTTTTTTTTTTTTTTTTTTTTTGTAGACAGGGTTTCACTCCATACAC	4020
QY	4021	CAGGCTGAGTGCAGATGGGGCAATCTTGGCTCACTGTAACTGTGACCTGCCCTCCAGGTCAA	4080
Dp	4021	CAGGCTGAGTGCAGATGGGGCAATCTTGGCTCACTGTAACTGTGACCTGCCCTCCAGGTCAA	4080
QY	4081	GCGATTCCTGCTTAAACCTCTGATAGCTGGAATTTATAGGCACACACCCACGACGCT	4140

Dd	4081	GGCATATTCCTCCGCTTACCTCCGAGAGTGGAAATTATAGGACACACACCACACGCT	4140
Qy	4141	GGCAATATTTTTTTTTTTTTTCTGTAATTTTTTACTAGAGACAGGGTTTCATCATGTTGGCAG	4200
Dd	4141	GGCAATATTTTTTTTTTTTTTCTGTAATTTTTTACTAGAGACAGGGTTTCATCATGTTGGCAG	4200
Qy	4201	GCTGGTCTTGAACCCCTGACTCAAGTAGATCCACCACCTCGGCTCCCAAGTGTGGG	4260
Dd	4201	GCTGGTCTTGAACCCCTGACTCAAGTAGATCCACCACCTCGGCTCCCAAGTGTGGG	4260
Qy	4261	ATTACAGGTGTACAGCCACATGACACAGGCCACATGTAATTTTTTAAATATTTTTTTA	4320
Dd	4261	ATTACAGGTGTACAGCCACATGACACAGGCCACATGTAATTTTTTAAATATTTTTTTA	4320
Qy	4321	ATTAAATGTTTATCTTAAGGCCAGTAGAGATCGGCTGTATATCCAGCACTTTGA	4380
Dd	4321	ATTAAATGTTTATCTTAAGGCCAGTAGAGATCGGCTGTATATCCAGCACTTTGA	4380
Qy	4381	GGGGCCAAAGTGTGGGGATCACTTAGGCTGGAGTTCCAGCTGGGCAACATAGTAGAC	4440
Dd	4381	GGGGCCAAAGTGTGGGGATCACTTAGGCTGGAGTTCCAGCTGGGCAACATAGTAGAC	4440
Qy	4441	CCCGTCTCTACCAAAAATTTAAAAAATTAGCTGGAGTGTGGCATTTGCTGTGTGCC	4500
Dd	4441	CCCGTCTCTACCAAAAATTTAAAAAATTAGCTGGAGTGTGGCATTTGCTGTGTGCC	4500
Qy	4501	AGCTACTTGGGAAGCTGAGGTGTGGGGATGGCTGAAGCTGTGAAGGTGAGAGCTGCAGT	4560
Dd	4501	AGCTACTTGGGAAGCTGAGGTGTGGGGATGGCTGAAGCTGTGAAGGTGAGAGCTGCAGT	4560
Qy	4561	AGCTATGATCACACACACTGCACTTCAAGCTGAGTACAGAGCTATCTCAAAAGCAAA	4620
Dd	4561	AGCTATGATCACACACACTGCACTTCAAGCTGAGTACAGAGCTATCTCAAAAGCAAA	4620
Qy	4621	ATAATGTTTATCTTAAGCGTAAAGTATATATACAGAAATATATGATAGATTTTAAATTGA	4680
Dd	4621	ATAATGTTTATCTTAAGCGTAAAGTATATATACAGAAATATATGATAGATTTTAAATTGA	4680
Qy	4681	AAAAGCATTAATGATTTACATGGATTGTAAATATCAAAATACATGAAATCTTGTTCTT	4740
Dd	4681	AAAAGCATTAATGATTTACATGGATTGTAAATATCAAAATACATGAAATCTTGTTCTT	4740
Qy	4741	AATAATGCTAGCAACAGCACATTTGGTTTTTACTAGGGCACCAGGTACTTTAAAAA	4800
Dd	4741	AATAATGCTAGCAACAGCACATTTGGTTTTTACTAGGGCACCAGGTACTTTAAAAA	4800
Qy	4801	AGTTAAGGCCACGCCACAGGGGCTCACACCTGTAATCCAGACATTTGGGAGGCCAAGCA	4860
Dd	4801	AGTTAAGGCCACGCCACAGGGGCTCACACCTGTAATCCAGACATTTGGGAGGCCAAGCA	4860
Qy	4861	GGAGGATACATTGAGGCCAGAGATTGGAGCTGAGCAACTATAGGAGATCTGTCTTG	4920
Dd	4861	GGAGGATACATTGAGGCCAGAGATTGGAGCTGAGCAACTATAGGAGATCTGTCTTG	4920
Qy	4921	TCTCTATAAAAAATTAAAAAATTGGCTAGGCCCTTTGGCTTACACCCGTAATCCAGCAC	4980
Dd	4921	TCTCTATAAAAAATTAAAAAATTGGCTAGGCCCTTTGGCTTACACCCGTAATCCAGCAC	4980
Qy	4981	TTTGGGAGGCCGAGGCGGCT 5000	
Dd	4981	TTTGGGAGGCCGAGGCGGCT 5000	
RESULT 2			
ABAI4559			
ID	ABAI4559 standard; DNA; 16650 BP.		
xx			
xx	ABAI4559;		
xx			
xx	23-JAN-2002 (first entry)		
DE	Human nervous system related polynucleotide seq ID NO 6890.		

1777 AAAAAGATCAGACTGAGAGGCTGCTCAACCCAGGCGAGAGGGTTTATACCAACCCCTCC 1836
1778 TCCACCCCGAGAGAGAGCCCTGAGAGCCAGGCTTTGCTGCTGCCCCCTCAGGCGCTG 3771
1837 TCCACCCCGAGAGAGAGCCCTGAGAGCCAGGCTTTGCTGCTGCCCCCTCAGGCGCTG 1896
3772 TTTACTGAATCCACACAGACTCATAGAGCCACATATAGATTATTAATAAGAGAGAGAG 3831
1897 TTTACTGAATCCACACAGACTCATAGAGCCACATATAGATTATTAATAAGAGAGAGAG 1956
3832 AGAGAGAGAGAGAGAGATGAGTCTCATCTGTGTCAGAGCGCTGCTCAGACTCTAG 3891
1957 AGAGAGAGAGAGAGAGATGAGTCTCATCTGTGTCAGAGCGCTGCTCAGACTCTAG 2016
3892 CTCAGCAATCCCGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3951
2017 CTCAGCAATCCCGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2076
3952 CACTTGACCAACACATGATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4011
2077 CACTTGACCAACACATGATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2136
4012 TCCATCACCAGCTGAGTGCAGTGGGGCAATCTTGCTCAGTGTATCTGCTGCTGCTGCTG 4071
2137 TCCATCACCAGCTGAGTGCAGTGGGGCAATCTTGCTCAGTGTATCTGCTGCTGCTGCTG 2196
4072 CAGGTGAGAGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4131
2197 CAGGTGAGAGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256
4132 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4191
2257 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2316
4192 GTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4251
2317 GTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
4252 AGTGTGAGGATTCAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4311
2377 AGTGTGAGGATTCAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436
4312 TATTTTAAATTAATAATGTTATCTAAGGCAAGTAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 4371
2437 TATTTTAAATTAATAATGTTATCTAAGGCAAGTAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2496
4372 GCACCTTGAAGGCGCAAGGCTGCGGGATCATTGAGCTGGAGTTCAAGCTGGGCAACA 4431
2497 GCACCTTGAAGGCGCAAGGCTGCGGGATCATTGAGCTGGAGTTCAAGCTGGGCAACA 2556
4432 TAGTGAAGCCCGCTCTACCAAAAATTTAAAAAATTTAGAGTGGAGTGGGCTGCTGCTGCTG 4491
2557 TAGTGAAGCCCGCTCTCTACCAAAAATTTAAAAAATTTAGAGTGGAGTGGGCTGCTGCTGCTG 2616
4492 TGTGTGCCAGCTACTGGGAACTGAGGTGTGGGATGCTGAAAGCTGTGAGGTGAG 4551
2617 TGTGTGCCAGCTACTGGGAACTGAGGTGTGGGATGCTGAAAGCTGTGAGGTGAG 2676
4552 GCTGAGAGAGTATGATACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4611
2677 GCTGAGAGTATGATACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2736
4612 GCAAAACAAATATGTTATCTAAGGCTAAGTATATACAGATATATATATATATATATATATAT 4671
2737 GCAAAACAAATATGTTATCTAAGGCTAAGTATATATATATATATATATATATATATATATAT 2796
4672 TTAATTTGAAAAAGCTTAT 4731
2797 TTAATTTGAAAAAGCTTAT 2856
4732 TGTGTTCTTAATTAATGCTAGCAACAGGCAATTTGTTTCTAGTGGGCAAGGCTGAC 4791

2857 TGTGTTCTTAATTAATGCTAGCAACAGGCAATTTGTTTCTAGTGGGCAAGGCTGAC 2916
4792 TTTAAAAAAGTTAGGCCAGCCACAGGAGGCTCAGACTGTATATCCAGACTTTGGAG 4851
2917 TTTAAAAAAGTTAGGCCAGCCACAGGAGGCTCAGACTGTATATCCAGACTTTGGAG 2976
4852 GCCAAGCAGAGAGATCATTGAGCCAGAGAGTTTGGAGCTTGGAGCAACATAGGAGATC 4911
2977 GCCAAGCAGAGAGATCATTGAGCCAGAGAGTTTGGAGCTTGGAGCAACATAGGAGATC 3036
4912 CTGATCTGCTCTATTAATAAATTTAAATAATTTGCTAGGCTTGTGCTTACACCCGTAA 4971
3037 CTGATCTGCTCTATTAATAAATTTAAATAATTTGCTAGGCTTGTGCTTACACCCGTAA 3096
4972 TCCAGCACTTTGGAGGCCGAGCGGGGT 5000
3097 TCCAGCACTTTGGAGGCCGAGCGGGGT 3125

RESULT 3
ABAI4557
ID ABAI4557 standard; DNA; 18530 BP.
XX
AC ABAI4557;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6888.
XX
KW Human; neotropic; neuroprotective; cytosolic; dermatological; vituicide;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antischilling; antianemic; antiallergic; cancer;
KW antileukemic; hepatocytic; cerebroprotective; antineoplastic;
KW antiallergic; antidiabetic; antiliver; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX
PN W0200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0198874.
PR 17-MAR-2000; 2000US-0198076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0218880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227109.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231958.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0251997.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system
cancers and metastases

Disclosure: SEQ ID NO 6888; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABAI1004-ABA21534) and proteins
(ABAI478-ABAI8001) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful
in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
and ovarian cancer and other cancers of the adrenal gland, bone, bone
marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
(b) immune disorders e.g. Addison's disease, allergies, autoimmune
haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
disease, multiple sclerosis, rheumatoid arthritis and ulcerative
colitis; (c) cardiovascular disorders such as myocardial ischaemia;
(d) wound healing; (e) neurological diseases e.g. cerebral ataxia and
epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
and parasitic infections.

Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;

Query Match 55.2%; Score 2758.2; DB 22; Length 18530;
Best Local Similarity 98.4%; Pred No. 0;
Matches 2803; Conservative 0; Mismatches 38; Indels 8; Gaps 2;

[illegible]

Db	1365	GGGCGCTGTGATCCCTTCAAAACCCCAATGGTGTCCCTCTACACGAGACCCAGCCAC	1424
QY	3292	CTGCTGACGGCCACTGCTGAGCCATGACATCACTTACACAAGCCAGTGGCTAATGCCC	3351
Db	1425	CTGCTGAGGCCACTGCTGAGCCATGACATCACTTACACAAGCCAGTGGCTAATGCCC	1484
QY	3352	GCTTAAAGCTCTTCTCCCGCTGCTGCTGCTGGCGGGGGGAGATTCACAGCTGCTAT	3411
Db	1485	GCTTAAAGCTCTTCTCCCGCTGCTGCTGCTGGCGGGGGGAGATTCACAGCTGCTAT	1544
QY	3412	ATGGCGAGTTCCTAATCTTCCTGCTGCTGCTTACATCATCCGCTTATTTATAGTAA	3471
Db	1545	ATGGCGAGTTCCTAATCTTCCTGCTGCTGCTTACATCATCCGCTTATTTATAGTAA	1604
QY	3472	GCTGGCAGGCTGGCGGGGGGGGCTTGGGAAGATGGCTGGGGCTGGAGCTGGAGC	3531
Db	1605	GCTGGCAGGCTGGCGGGGGGGGCTTGGGAAGATGGCTGGG-----GCTGGAGC	1657
QY	3532	TCCGAGGGGGCTCCAGCACAGCTCAGAGGCCAGTGCACAGTCCATACACATPAAGT	3591
Db	1658	TCCGAGGGGGCTCCAGCACAGCTCAGAGGCCAGTGCATACATACACATPAAGT	1717
QY	3592	GGGCTCTGACACAGCTCTTGGGCACTGAGCTAGGCTCGGGCGTGGGGGGTGGGAGAG	3651
Db	1718	GAGCTCTGACACAGCTCTTGGGCACTGAGCTAGGCTG--GGCTTGGGGCTGGGAGAG	1776
QY	3652	TAAAGAGTCACTGAGAGGCTGCTCAAGCCAGGAGAGGGTTTAAAGCACCTTCC	3711
Db	1777	AAAGAGTCACTGAGAGGCTGCTCAAGCCAGGAGAGGGTTTAAAGCACCTTCC	1836
QY	3712	TCCAACCCGAGAGAGACCCCTGGAGCCAGGCTTTGTCTGGCCCACTGACGGCTGT	3771
Db	1837	TCCAACCCGAGAGAGACCCCTGGAGCCAGGCTTTGTCTGGCCCACTGACGGCTGT	1896
QY	3772	TTTACTGAATCCGACACAGACTCTATAGGCCACATATGATTTAAAAAGAGAGAGAG	3831
Db	1897	TTTACTGAATCCGACACAGACTCTATAGGCCACATATGATTTAAAAAGAGAGAGAG	1956
QY	3832	AGAGAGAGAGAGAGATGAGTCACTGTGTGTCCAGGCTGCTGCAATCTCTAGG	3891
Db	1957	AGAGAGAGAGAGAGATGAGTCACTGTGTGTCCAGGCTGCTGCAATCTCTAGG	2016
QY	3892	CTCAAGCAATCCCTGCTTAGCTGCCAAGGGGCTGGGATTACAGGTGTGACTG	3951
Db	2017	CTCAAGCAATCCCTGCTTAGCTGCCAAGGGGCTGGGATTACAGGTGTGACTG	2076
QY	3952	CACCTGACCAACACATGTACTTTTTTTTTTTTTTTTTTTTGAGACAGGTTTCAC	4011
Db	2077	CACCTGACCAACACATGTACTTTTTTTTTTTTTTTTTTTTGAGACAGGTTTCAC	2136
QY	4012	TCATATCACCAAGCTGAGTGCAGTGGGGGGAATCTGGGCTCACTGTAACTCGCTCC	4071
Db	2137	TCATATCACCAAGCTGAGTGCAGTGGGGGGAATCTGGGCTCACTGTAACTCGCTCC	2196
QY	4072	CAGGTCAAGCAATCTCTGCTTAGCTCCTGAGTAGCTGAATTATAGCACACACC	4131
Db	2197	CAGGTCAAGCAATCTCTGCTTAGCTCCTGAGTAGCTGAATTATAGCACACACC	2256
QY	4132	ACCAGCCTGGCTAATTTTTTTTTTTTTTTCTGTATTTTATAGAGACAGGTTTCATCAT	4191
Db	2257	ACCAGCCTGGCTAATTTTTTTTTTTTTTTCTGTATTTTATAGAGACAGGTTTCATCAT	2316
QY	4192	GTTGGGCAGGCTGCTTGAACCCCTGACCTCAAGTGAATCCACCCACTCGGCTCCAA	4251
Db	2317	GTTGGGCAGGCTGCTTGAACCCCTGACCTCAAGTGAATCCACCCACTCGGCTCCAA	2376
QY	4252	AGTGTGGATTACAGGTGTACAGCCACATGACACAGCCACATGTAATTTTAAAT	4311
Db	2377	AGTGTGGATTACAGGTGTACAGCCACATGACACAGCCACATGTAATTTTAAAT	2436
QY	4312	TATTTTATATTAATGTTTATCTTAAGGCTGATGACAGTGAAGCTGCTGTATTCCA	4371

Db 2437 TATTTTTAAATTAATGTTTATCTAAGCCAGTAGACTCGGCTGTGAATCCCA 2496
OY 4372 GCACCTTTGAGGGGCCAAGGTGCGGGATCATCTTGACCTTGGAGTTGAGCGTGACACA 4431
Db 2497 GCACCTTTGAGGGGCCAAGGTGCGGGATCATCTTGAGCTTGGAGTTGAGCGTGACACA 2556
OY 4432 TACTGAGACCCCTCTCTACCAAAAATTTAAAAAATTTAGCTGGAGTGGGCACTTTGCC 4491
Db 2557 TACTGAGACCCCTCTCTACCAAAAATTTAAAAAATTTAGCTGGAGTGGGCACTTTGCC 2616
OY 4492 TGTGTGCCAGCTACTTGGAGACTGAGTGTGGGATGCTGAAAGCTGTGAGCTGAG 4551
Db 2617 TGTGTGCCAGCTACTTGGAGACTGAGTGTGGGATGCTGAAAGCTGTGAGCTGAG 2676
OY 4552 GCTGAGTGAAGCTATGATACACACACTGACCTTACAGCTAGTACAGGCTATCTCAAA 4611
Db 2677 GCTGAGTGAAGCTATGATACACACACTGACCTTACAGCTAGTACAGGCTATCTCAAA 2736
OY 4612 GCAACCAAAATATGTTATCTTAAACGGTAAAGTATATCAAGAAATATATGATGAT 4671
Db 2737 GCAACCAAAATATGTTATCTTAAACAAATATATATATATATATATATATATGATG 2796
OY 4672 TTAATTTGAAAAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATG 4731
Db 2797 TTAATTTGAAAAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATG 2856
OY 4732 TGTGTCTTAAATATGCTAGCAACAGGCAATTTGTTTACTAGGGCCCAAGGTAC 4791
Db 2857 TGTGTCTTAAATATGCTAGCAACAGGCAATTTGTTTACTAGGGCCCAAGGTAC 2916
OY 4792 TTTAAAAAAGTTAGGGCCAGCCAGGAGGCTCACCTGTAAATCCAGCACTTTGGAG 4851
Db 2917 TTTAAAAAAGTTAGGGCCAGCCAGGAGGCTCACCTGTAAATCCAGCACTTTGGAG 2976
OY 4852 GCCAAGCAGAGAGATCACTTGAAGCCAGAGATTAGAGCTGACCAATAGGAGATC 4911
Db 2977 GCCAAGCAGAGAGATCACTTGAAGCCAGAGATTAGAGCTGACCAATAGGAGATC 3036
OY 4912 CTGATCTTGTCTCTATTAATAAATTTAAATTTGGCTTGGCTTACACCCGTAA 4971
Db 3037 CTGATCTTGTCTCTATTAATAAATTTAAATTTGGCTTGGCTTACACCCGTAA 3096
OY 4972 TCCAGAGCTTTGGAGCCGAGGCGGCT 5000
Db 3097 TCCAGAGCTTTGGAGCCGAGGCGGCT 3125

RESULT 4
ID ABA14558 standard; DNA; 18537 BP.
XX ABA14558;
AC ABA14558;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6889.
XX
KW Human; neotropic; neuroprotective; cytosolic; dermatological; vitruclae;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; antiviral;
KW antiparkinsonian; antidiabetic; antianemic; antirheumatic; cancer;
KW antineoplastic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234697.
PR 25-SEP-2000; 2000US-0234996.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM-nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:10:38 ; Search time 317.898 Seconds
(without alignments)
10693.617 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429
Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1244.2	51.2	1263	9	US-09-746-783-3
2	350.4	14.4	1326	9	US-10-071-766-10
3	189.2	7.2	1292	9	US-10-198-846-11070
4	175	7.8	1198	10	US-09-880-107-2174
5	126.8	5.2	1350	10	US-09-768-826-16
6	116	4.8	853	9	US-10-198-846-11246
7	104	4.3	462	9	US-10-198-846-2561
8	102	4.0	615	9	US-10-198-846-7843
9	98	4.0	507	9	US-09-764-872-195
10	90	3.7	684973	10	US-09-263-959-1
11	89	3.7	31994	9	US-09-764-904-71
12	89	3.7	31994	9	US-10-091-548-71
13	89	3.7	31994	9	US-10-074-095-599
14	89	3.7	31994	10	US-09-764-860-599
15	88	3.6	30568	10	US-09-764-877-3851
16	87.8	3.6	133893	9	US-10-161-510-1
17	87.4	3.6	4863	10	US-09-764-877-2903
18	87.4	3.6	25309	9	US-10-291-737-3
19	86.4	3.6	2625	9	US-09-764-891-8802

20	86.2	3.5	289	9	US-10-115-278-4	Sequence 4, Appl1
21	86.2	3.5	8663	9	US-10-079-854-397	Sequence 397, App
22	86.2	3.5	8663	10	US-09-764-878-397	Sequence 397, App
23	86.2	3.5	75899	10	US-09-854-883-243	Sequence 211, App
24	86	3.5	10894	9	US-10-091-504-2111	Sequence 2111, App
25	86	3.5	10894	10	US-09-764-869-2111	Sequence 5805, App
26	85.8	3.5	219	9	US-09-764-891-5805	Sequence 5805, App
27	85.8	3.5	219	9	US-09-764-891-5806	Sequence 32677, A
28	85.8	3.5	484	9	US-09-918-995-32677	Sequence 8703, App
29	85.8	3.5	8018	9	US-09-764-891-8703	Sequence 3468, App
30	85.8	3.5	16337	10	US-09-764-877-3468	Sequence 1605, App
31	85.8	3.5	32195	9	US-10-091-504-1605	Sequence 611, App
32	85.8	3.5	32195	9	US-10-125-540-611	Sequence 611, App
33	85.8	3.5	32195	10	US-09-764-870-611	Sequence 617, App
34	85.8	3.5	32195	10	US-09-764-870-617	Sequence 1605, App
35	85.8	3.5	32195	10	US-09-764-869-1605	Sequence 6186, App
36	85.8	3.5	32195	10	US-09-764-891-6187	Sequence 6187, App
37	85.6	3.5	261	9	US-09-764-891-6187	Sequence 6188, App
38	85.6	3.5	261	9	US-09-764-891-6188	Sequence 2901, App
39	85.6	3.5	261	9	US-10-198-846-2901	Sequence 260, App
40	85	3.5	477	9	US-09-844-653-5	Sequence 6304, App
41	85	3.5	402850	9	US-09-844-653-5	Sequence 1373, App
42	84.8	3.5	2581	9	US-09-822-846-260	
43	84.8	3.5	32191	9	US-09-764-891-6304	
44	84.6	3.5	184	9	US-09-764-868-1373	
45	84.6	3.5	288	9	US-10-115-278-2	

ALIGNMENTS

RESULT 1
US-09-746-783-3
Sequence 3, Appl1
Publication No. US2003044935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Racle, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtner, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match 51.2%; Score 1244.2; DB 9; Length 1263;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

1181 GACAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAAGCAACTGGATTGTC 1240
Db 1 GAGAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAAGCAACTGGATTGTC 60
QY 1241 GACAGCAATTTGCAGCTGTCCTGCTGGTGGATGATGATGACAGCACTGCTCGG 1300
Db 61 GACAGCAATTTGCAGCTGTCCTGCTGGTGGATGATGATGACAGCACTGCTCGG 120
QY 1301 ATGAGAGCGGACATGCTAGTGAATTAAGCCGAGCCAGACGCCCTTACAGAGCTGCTCC 1360
Db 121 ATGAGAGCGGACATGCTAGTGAATTAAGCCGAGCCAGACGCCCTTACAGAGCTGCTCC 180
QY 1361 GCCAGATTCCGTGAGCCCTCTTATAGGCTCCACCTTCAACATCAGCCTGAACAAAGAG 1420
Db 181 GCCAGATTCCGTGAGCCCTCTTATAGGCTCCACCTTCAACATCAGCCTGAACAAAGAG 240
QY 1421 GAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCAGCGTGGCATCTTGGC 1480
Db 241 GAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCAGCGTGGCATCTTGGC 300
QY 1481 CGCTTCTTAGGCTGCACTGCCATGATCACCATCTCCCAAGGGCAAACTCAAGAGCCAAA 1540
Db 301 CGCTTCTTAGGCTGCACTGCCATGATCACCATCTCCCAAGGGCAAACTCAAGAGCCAAA 360
QY 1541 CTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGGGCTGCCCCAAAACACAGAGCA 1600
Db 361 CTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGGGCTGCCCCAAAACACAGAGCA 420
QY 1601 GCCAAACAGAACGTTAGGGGCCAGAGAACAAAGAGCCCTGGAAGCTTAAAGCTGTGAGC 1660
Db 421 GCCAAACAGAACGTTAGGGGCCAGAGAACAAAGAGCCCTGGAAGCTTAAAGCTGTGAGC 480
QY 1661 GCCCTCAAGTGTGCCCCACTGTATCAGAGGCCACTACTACAGTGGCCCCACAGAGGCC 1720
Db 481 GCCCTCAAGTGTGCCCCACTGTATCAGAGGCCACTACTACAGTGGCCCCACAGAGGCC 540
QY 1721 CTCAGGCCCACTCCCATGTTCTTCCGCCCTAGAACATCAGCGCTCAAGCTTCAAGT 1780
Db 541 CTCAGGCCCACTCCCATGTTCTTCCGCCCTAGAACATCAGCGCGCTCAAGCTTCAAGT 600
QY 1781 GTTCACAGGCAATAGACACCAAGACAAAGACCTTAAAGACTGTGAGTTCTGGGCCAAAGAA 1840
Db 601 GTTCACAGGCAATAGACACCAAGACAAAGACCTTAAAGACTGTGAGTTCTGGGCCAAAGAA 660
QY 1841 AGTTTGAATGTCTCTCAAGAGCGATGGGGCCCTTATGAGAGACCCCAAGAAATCTCAA 1900
Db 661 AGTTTGAATGTCTCTCAAGAGCGATGGGGCCCTTATGAGAGACCCCAAGAAATCTCAA 720
QY 1901 GTGAGAGAGAAACTGTGAGATTAACTGACGATATGCGAGATATCCCGAATAATCAGC 1960
Db 721 GTGAGAGAGAAACTGTGAGATTAACTGACGATATGCGAGATATCCCGAATAATCAGC 780
QY 1961 CTCAGAAACCTTTGGAACATATCAGCAACATCAGCTCAGCTCAAAAGATCTCAGG 2020
Db 781 CTCAGAAACCTTTGGAACATATCAGCAACATCAGCTCAGCTCAAAAGATCTCAGG 840
QY 2021 GATCCTTATGGGCTTGAAGAAACAGGAGATGAAGCAATCTTAACTGCTGCTCTAATG 2080
Db 841 GATCCTTATGGGCTTGAAGAAACAGGAGATGAAGCAATCTTAACTGCTGCTCTAATG 900
QY 2081 GGATGCTTGGCCAGCGATGCTCAGCTGTGTGACACGACGAGACATGATCCAGTCA 2140
Db 2141 GGATGCTTGGCCAGCGATGCTCAGCTGTGTGACACGACGAGACATGATCCAGTCA 2200
QY 2141 CAGCCATACAGCTGCTCCACACTGGAAGAAAGCGTCTTACAAAGCCTGAATGTTA 2200
Db 961 CAGCCATACAGCTGCTCCACACTGGAAGAAAGCGTCTTACAAAGCCTGAATGTTA 1020
QY 2201 GCTTAATGATTAATAATCCAGACTACTTCAAGCTTTAATGCTTTTATCATAAACT 2260
Db 1021 GCTTAATGATTAATAATCCAGACTACTTCAAGCTTTAATGCTTTTATCATAAACT 1080
QY 2261 GTGAAAGCTAGACTGACCAATTTGAACATTTAACTCAGACTCTGAGATTCAGAGCGGGA 2320
Db 1081 GTGAAAGCTAGACTGACCAATTTGAACATTTAACTCAGACTCTGAGATTCAGAGCGGGA 1140
QY 2321 ACCCTTAGTCTTATCTGAAATCCAGAGCCAGCCACTTATAGTATCTGCCCCAACTAATGA 2380
Db 1141 ACCCTTAGTCTTATCTGAAATCCAGAGCCAGCCACTTATAGTATCTGCCCCAACTAATGA 1200
QY 2381 GTTTAAATATCAATATCTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429
Db 1201 GTTTAAATATCAATATCTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1249
```

RESULT 2

US-10-071-766-10/c
Sequence 10, Application US/10071766
Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
FILE REFERENCE: PA-0043 US
CURRENT APPLICATION NUMBER: US/10/071,766
CURRENT FILING DATE: 2002-02-07
NUMBER OF SEQ. ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 10
LENGTH: 1326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incycle ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 14.4%; Score 350.4; DB 9; Length 1326;
Best Local Similarity 99.7%; Pred. No. 5; Se-91;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

2053 AGCAATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGCTCAGCTGTGT 2112
Db 1326 AGCAATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGCTCAGCTGTGT 1267
QY 2113 GTACACGAGAGAGACAGTATCAGTACAGCCATACAGCTGTCCACACTGAAAGACGTG 2172
Db 1266 GTACACGAGAGAGACAGTATCAGTACAGCCATACAGCTGTCCACACTGAAAGACGTG 1207
QY 2173 TCCTAACAACCTGGAATCAAAATGCTTAATAGTAAATTAATTAATTAATTAATTAAT 2232
Db 1206 TCCTAACAACCTGGAATCAAAATGCTTAATAGTAAATTAATTAATTAATTAATTAAT 1147
QY 2233 CCTTAATGCTTTTATCATTAATAAATGTAAGAGCTAGTGAACATTGGAACAATT 2292
Db 1146 CCTTAATGCTTTTATCATTAATAAATGTAAGAGCTAGTGAACATTGGAACAATT 1087
QY 2293 AACTCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGAATCCAAAGACGCA 2352
Db 1086 AACTCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGAATCCAAAGACGCA 1027
QY 2353 CACCTTAGTATCTGCCAAACTAATGAGTTTAAATTAATTAATTAATTAATTAAT 2404
Db 1026 CACCTTAGTATCTGCCAAACTAATGAGTTTAAATTAATTAATTAATTAATTAAT 975
```

RESULT 3
US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198, 846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 7.8%; Score 189.2; DB 9; Length 1292;
Best Local Similarity 62.4%; Pred. No. 3.2e-44;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 349 TGGGCTTCTACGTGAGCGTGTGTCGACCCGCTGTGAGAACAGTACGAGAACCTGCGCT 408
DB 318 TAGGGTTTATGTACTCTGTAAGAACCGATGCTGAGAACCAAGTTGATTTGCGCT 377
QY 409 GGGCCGACCGCTCTAGACCTGCTGGGCTTGTGAGAGGCAAGACGACGACGAC 468
DB 378 GGGCCGACGAGCGTAACTGCTCACTCTAGACAGTGTTCACGAGACGACGACGACG 437
QY 469 GGGCTGCGGCGGACGCTATCCGCTAGCCAGCCAGCTGGGCAAGCTGCTCATCTCGCA 528
DB 438 GCGTCTTGAAGAGAGCGATGCGGTAGCTCACTCTGCTCATCTCTGCT 497
QY 529 GGGTACGACCGCAGTCTCAAGCGCTTCCCAAGGCGCCAGACCTGGTGGCAAGAGCT 588
DB 498 CGGTGAGCACTGCTGTGTACAAAGATTTCCCAAGTGAACGACGAGTGTGAAGAGCT 557
QY 589 TTTATGCTCGGCGAGAACAGCAAGTGTGAGAACTGAGCACTACCAACAACATGTTCT 648
DB 558 TTTATGCAACAGATGAAAGAAATTTTCAACCACTCAAGTCTCTCATCTGAATATT 617
QY 649 GGGTGGCTGGGTGTGTGTTGCAACCTGTCAATGAAGCGGTGGCTTGAAGCTCAATCC 708
DB 618 GGGTTCATCATCTGCTGTTGAAATCTTGCAACTGAAGCCGGAATGAAGTGAATCA 677
QY 709 GGGACCTATCTGCTCCAGAGCTGCTGGAAGAGATGAACACCTGCGTCACTGCTG 768
DB 678 GAGACAGTGTGATCTGCAATCTGATGACTGAATGAATGATCATACCGCTTGGTGCA 737
QY 769 GACACCTGTATGCTTACGATGATTTATCCCACTGATGTATACACAGGTG 822
DB 738 GCCTTTATTCGGTTATGACTGGGTTCGCGTGGTTACACCCAGTAG 791

RESULT 4
US-09-880-107-2174/C
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darcil T.
APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174

Query Match 7.2%; Score 175; DB 10; Length 1198;
Best Local Similarity 100.0%; Pred. No. 4e-40;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2230 CAGCCTTATGCTTTTATTCATTAAGCTGGAAGCTAGACAGCACTTGGAAACA 2289
DB 1198 CAGCCTTATGCTTTTATTCATTAAGCTGGAAGCTAGACAGCACTTGGAAACA 1139
QY 2290 TTTAAGCAGACCTGATGATGAGTGGGAAACCTTACTTCTATCTGAATCCAGACG 2349
DB 1138 TTTAAGCAGACCTGATGATGAGTGGGAAACCTTACTTCTATCTGAATCCAGACG 1079
QY 2350 CCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2404
DB 1078 CCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024

RESULT 5
US-09-768-826-16
Sequence 16, Application US/09768826
Patent No. US20020012966A1
GENERAL INFORMATION:
APPLICANT: Shi et al.
TITLE OF INVENTION: 18 human secreted proteins
FILE REFERENCE: PFS12P1
CURRENT APPLICATION NUMBER: US/09/768,826
CURRENT FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/US00/22350
PRIOR FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: 60/148,759
PRIOR FILING DATE: 1999-08-16
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 1350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c

```

1 NAME/KEY: SITE
2 LOCATION: (1209)
3 OTHER INFORMATION: n equals a,t,g, or c
4 NAME/KEY: SITE
5 LOCATION: (1229)
6 OTHER INFORMATION: n equals a,t,g, or c
7 NAME/KEY: SITE
8 LOCATION: (1266)
9 OTHER INFORMATION: n equals a,t,g, or c
10 NAME/KEY: SITE
11 LOCATION: (1285)
12 OTHER INFORMATION: n equals a,t,g, or c
13 NAME/KEY: SITE
14 LOCATION: (1287)
15 OTHER INFORMATION: n equals a,t,g, or c
16 NAME/KEY: SITE
17 LOCATION: (1280)
18 OTHER INFORMATION: n equals a,t,g, or c
19 NAME/KEY: SITE
20 LOCATION: (1295)
21 OTHER INFORMATION: n equals a,t,g, or c
22 NAME/KEY: SITE
23 LOCATION: (1305)
24 OTHER INFORMATION: n equals a,t,g, or c
25 NAME/KEY: SITE
26 LOCATION: (1324)
27 OTHER INFORMATION: n equals a,t,g, or c
28 NAME/KEY: SITE
29 LOCATION: (1339)
30 OTHER INFORMATION: n equals a,t,g, or c
31 NAME/KEY: SITE
32 LOCATION: (1341)
33 OTHER INFORMATION: n equals a,t,g, or c
34 NAME/KEY: SITE
35 LOCATION: (1343)
36 OTHER INFORMATION: n equals a,t,g, or c
37 NAME/KEY: SITE
38 LOCATION: (1345)
39 OTHER INFORMATION: n equals a,t,g, or c
40 NAME/KEY: SITE
41 LOCATION: (1348)
42 OTHER INFORMATION: n equals a,t,g, or c
43
44 JS-09-768-826-16

```

	Query Match	5.2%	Score 126.8	DB 10	Length 1350	
	Best Local Similarity	69.9%	Pred. No. 4,1e-26			
	Matches 186	Conservative	0	Mismatches 77	Indels 3	Gaps 1
QY	1174	GGTGGCAGCAGCATCATCAACCCCTTTGGAGAGGATGATGTGAATTTTGAGACCACACTG	1233			
Db	436	GGTGGCTGACAGATCATCAACCCATTGGTGGAGATGATGACGACTTTGAAACAATCA	495			
QY	1234	GATTGTGCACAGGAATTTTCAGAGTCTCCCTGTTCGCTGTGGATGAGTGCACCAAGACT	129			
Db	496	GCTATAACCGCAACTTCGAGGTCTCCCTGTTATCCGTGACGAAATGTACCAAACT	555			
QY	1294	GCCTCGATGAGCGGACATGTACTGTGAATAAGCCGACACGCCGCCCTTACAC---	135			
Db	556	TCCCCCGCCTAGAAGGACACACTTGGAGTAGAACAACGCCGACACCTTAACACT	615			
QY	1351	AGCTGCTTCCGCCAGTTCGCTGAGACCTCCTTTATGGAGCTCCACCTTCACATCAGCCT	1410			
Db	616	GGCCACGGCGCGCGAGTCTGTGCGGCCCTCATTTCTGGGCTCCACCTTCAACCTCGCAT	675			
QY	1411	GAACAAAGAGAGATGAGATTCAC	1436			
Db	676	GAGCGACGACCCTGACAGACCTTCG	701			

```

1 GENERAL INFORMATION:
2 APPLICANT: Lillie, James
3 APPLICANT: Xu, Yongyao
4 APPLICANT: Wang, Youzhen
5 APPLICANT: Steilmann, Kathleen
6 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
7 TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
8 TITLE OF INVENTION: THERAPY OF BREAST CANCER
9 FILE REFERENCE: MRI-049
10 CURRENT APPLICATION NUMBER: US/10/198,846
11 CURRENT FILING DATE: 2002-07-18
12 PRIOR APPLICATION NUMBER: 60/306,220
13 PRIOR FILING DATE: 2001-07-18
14 NUMBER OF SEQ. ID NOS: 14084
15 SOFTWARE: FastSeq for Windows Version 4.0
16 SEQ ID NO 11346
17 LENGTH: 853
18 TYPE: DNA
19 ORGANISM: Homo sapiens
20 FEATURE:
21 NAME/KEY: misc.feature
22 LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41
23 LOCATION: 43, 50, 51, 54, 57, 853
24 OTHER INFORMATION: n = A,T,C or G
25 US-10-198-846-11346

```

	Query Match	4.8%	Score 116	DB 9	Length 853
	Best Local Similarity	68.2%	Pred. No. 43e-23		
	Matches 161	Conservative 0	Mismatches 75	Indels 0	Gaps 0
OY	350	GGGCTTACGTGACGCTGTCGTGTGACCCGCTGTGGAAACAGTACAGAGAACCTGCGCTG	409		
Db	272	GGGGTTTATTTACTCTGTGTAGTAAACGAAGGTGGAAACCGTGTGTAATTTGGCCCTG	331		
OY	410	GCCGACCGCCTCATGACCTGTGTGTGGGCTTCTCGAAGCAAGAACGACAGACCGG	469		
Db	332	GCCAGACAGGCTAATGTTCTTCATCTTACGACAGTTCACGGAACGACGACGCGGG	391		
OY	470	GCTGCTGCGGCGACGCTCATCCGCTACGCCAACCTGGGCAACGTGCTCATCTGTGCGCA	529		
Db	392	CCTGCTTTGAAGAGAGCGCTGATGCGCTACGTCAATCTCACTCCCTGCATCTATTTTCGCTC	451		
OY	530	CGTAGACGACGACAGTCTACAAAGCGTTCCCCGAGCGCCGACGACCTGGTGCACGAG	585		
Db	452	GGTGAACCACTGCTGTGTACAAAAGATTTTCCCACTATGCAACCAAGCGTGGTTGAAGCAG	507		

```

RESULT 7
US-10-198-846-2561/c
: Sequence 2561, Application US/10198846
: Publication NO. US2003009997A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinhmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2561
: LENGTH: 462
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 390..462

```


REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 920010.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-8900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 684973 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-263-959-1

Query Match 3.7%; Score 89; DB 9; Length 684973;
Best Local Similarity 86.7%; Pred. No. 2.6e-14;

Matches 102; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1079 AGAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGC 1138

DB 118384 AAAATATCAGCCAGGCTGTGTGGTGCACACTGTATCCAGTACTCGGAGGCTGAGGC 118325

QY 1139 AGAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCAACCC 1198

DB 118324 AGAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCAACCC 118265

QY 1139 TT 1200

DB 118264 TT 118263

RESULT 11

US-09-764-904-71/c

Sequence 71, Application US/09764904

Patent No. US20020173454A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PAL122

CURRENT APPLICATION NUMBER: US/09/764,904

Prior Application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 137

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 71

LENGTH: 31994

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-904-71

Query Match 3.7%; Score 89; DB 9; Length 31994;

Best Local Similarity 86.7%; Pred. No. 2.6e-14;

Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1081 AAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140

DB 6484 AAATATGACGAGGTGTGTGGTGCATCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425

QY 1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193

DB 6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

RESULT 12

US-10-091-548-71/c

Sequence 71, Application US/10091548

Publication No. US20030049703A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PAL122C1

CURRENT APPLICATION NUMBER: US/10/091,548

Prior Application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 137

Prior Application removed - See File Wrapper or Palm
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 71
LENGTH: 31994
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-548-71

Query Match 3.7%; Score 89; DB 9; Length 31994;

Best Local Similarity 86.7%; Pred. No. 2.6e-14;

Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1081 AAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140

DB 6484 AAATATGACGAGGTGTGTGGTGCATCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425

QY 1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193

DB 6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

RESULT 13

US-10-074-095-599/c

Sequence 599, Application US/10074095

Publication No. US20030077704A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC008C1

CURRENT APPLICATION NUMBER: US/10/074,095

Prior Application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 137

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 71

LENGTH: 31994

TYPE: DNA

ORGANISM: Homo sapiens

US-10-074-095-599/c

Query Match 3.7%; Score 89; DB 9; Length 31994;

Best Local Similarity 86.7%; Pred. No. 2.6e-14;

Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1081 AAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140

DB 6484 AAATATGACGAGGTGTGTGGTGCATCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425

QY 1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193

DB 6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

QY 1081 AAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140

DB 6484 AAATATGACGAGGTGTGTGGTGCATCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425

QY 1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193

DB 6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

QY 1081 AAAGACCCAGGCGTGTGGCGCACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140

DB 6484 AAATATGACGAGGTGTGTGGTGCATCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425

QY 1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193

DB 6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

PRIOR APPLICATION NUMBER: 60/224, 518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224, 519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220, 964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241, 809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249, 299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236, 327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241, 785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244, 617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225, 268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236, 368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251, 856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251, 868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229, 344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234, 997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229, 513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231, 413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229, 509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236, 367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237, 039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236, 370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236, 802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237, 040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240, 960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239, 935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239, 937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241, 787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246, 474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246, 532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249, 216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226, 681

PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225, 759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225, 213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227, 182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225, 214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235, 836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230, 438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215, 135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225, 266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249, 218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249, 297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232, 400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231, 242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232, 081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232, 080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231, 414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231, 244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233, 064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233, 063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232, 397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232, 399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232, 401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241, 808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241, 826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241, 786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241, 221
PRIOR FILING DATE: 2000-10-20

```

QY 1079 AGAAGCAGCAGCGCGTGTGCGGCACACCTGTAAATCCAGCTACTCGGAGAGGTGAGGC 1138
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27997 AAAAAATTAGCAGACGCTGTGTGCGGCGCACCTGTAAATCCAGCAGCTCGGAGGCTGAGGC 28056
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1139 AGGAGAAATCGCTTGAACCCGGGAGGCGGAGGTGTGTGGCAGAGCAGCTCATCAACCCC 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28057 TGGAGAAATCGCTTGAACCCGGGAGGCGGAGGTGTGTGGCAGAGCAGCTCATCAAAATCG 28116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1199 TTTGGAGAGGAGTATGATGATTTT 1222
    ||| ||||| ||| ||| |||||
Db 28117 TTAGAGAGTACTTTTTTTTTTTT 28140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:29:22 ; Search time 488.416 Seconds

(without alignments)
11199.691 Million cell updates/sec

Title: US-09-622-964-4
Perfect score: 2429
Sequence: 1 caggaggtccacacagccta.....aaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2429	100.0	2429	20	AAZ21228 Human CGICE long f
2	2013	82.9	2229	20	AAZ21227 Human CGICE short
3	1244.2	51.2	1263	20	AAV99722 Human adult retina
4	658.2	27.1	1916	20	AAZ21229 Mouse CGICE cDNA s
5	642.2	26.4	16125	20	AAZ21226 Human CGICE genom
6	642.2	26.4	16650	22	ABAI4559 Human nervous syst
7	642.2	26.4	18530	22	ABAI4557 Human nervous syst
8	640.6	26.4	18537	22	ABAI4558 Human nervous syst
9	587.8	24.2	1717	24	ABL89697 Human polynucleoti

10	476.4	19.6	7108	22	ABAI4556 Human nervous syst
11	294.6	12.1	305	22	ABAI1094 Drosophila melanog
12	237	9.8	2861	23	ABL10793 EST clone Bk112.
13	234.8	9.7	424	20	AAV86976 Drosophila melanog
14	193	7.9	1608	23	ABL12609 Drosophila melanog
15	188.2	7.7	10760	23	ABL10792 Drosophila melanog
16	188.2	7.2	1198	24	ABK83853 Human cDNA differe
17	175	7.2	1198	24	ABN95676 Gene #2174 used to
18	175	7.2	1234	23	AA567682 DNA encoding novel
19	175	7.2	1279	23	AA585665 DNA encoding novel
20	175	7.2	1279	23	AA591587 DNA encoding novel
21	154.8	6.4	1526	23	AA572796 DNA encoding novel
22	150.4	6.2	3592	23	ABL12636 Drosophila melanog
23	146.8	6.0	1345	23	ABL12637 Drosophila melanog
24	126.8	5.2	1350	22	AA76848 Human secreted pro
25	120.4	5.0	2196	23	ABL16195 Drosophila melanog
26	115	4.7	2104	23	AA572795 DNA encoding novel
27	104.8	4.3	3785	23	ABL12608 Drosophila melanog
28	100.8	4.1	392	22	ABL11210 Human breast cance
29	100.8	4.1	401	22	ABL20109 Human breast cance
30	99.8	4.1	527	22	ABL17589 Human breast cance
31	98	4.0	507	22	AA539532 cDNA encoding nove
32	98	4.0	507	22	AAK88431 Human digestive sy
33	90	3.7	267156	24	ABL68560 Kidney cancer rela
34	89	3.7	514	22	AAH13186 Human cDNA clone (
35	89	3.7	31994	22	AA530619 DNA encoding novel
36	89	3.7	31994	22	AA528165 Genomic sequence #
37	88.8	3.7	10366	22	AAK71267 Human immune/haema
38	88.8	3.7	10366	22	AAK90439 Human digestive sy
39	88.8	3.7	26427	22	ABA20762 Human nervous syst
40	88.8	3.7	26427	22	ABA20763 Human nervous syst
41	88.2	3.6	8056	22	AAK79815 Human immune/haema
42	88	3.6	30568	22	ABL37486 Human musculoskele
43	87.8	3.6	139904	24	ABK83562 Human cDNA differe
44	87.6	3.6	19199	22	AAK70995 Human immune/haema
45	87.4	3.6	370	14	AAQ59541 Human brain Expres

ALIGNMENTS

RESULT 1	
ID	AAZ21228 standard; cDNA; 2429 BP.
AAZ21228	
AC	AAZ21228;
XX	
XX	
DT	22-NOV-1999 (first entry)
DE	Human CGICE long form cDNA sequence.
XX	
XX	
KW	CGICE; Best's macular dystrophy; mutation; diagnosis; detection;
KM	BMD; age-related macular dystrophy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	105..1412
FT	/*tag= a
FT	/product= "CGICE short form protein"
PN	W09943695-A1.
XX	
PD	02-SEP-1999.
XX	
XX	
PF	22-FEB-1999; 99WO-US03790.
XX	
PR	25-FEB-1998; 98US-0075941.
PR	18-DEC-1998; 98US-0112926.
XX	
PA	(MERI) MERCK & CO INC.
XX	(UYUP-) UNIV UPSALA.

Db	1801	AGACAAAAGCTTAAAGCTGTGAGTCTGTGGGCCAAGAAAAGTTTGAATTGCTCTCAGA	1860
Oy	1861	GAGCGATGGGGCCCTTGATGGAGACACCAGAAAGTATCTCAAGTGAGAGGAAAACTGTGCA	1920
Db	1861	GAGCGATGGGGCCCTTGATGGAGACACCAGAAAGTATCTCAAGTGAGAGGAAAACTGTGCA	1920
Oy	1921	GTTTAACCTTACCGGANTATGCCAGAGATCCCCGAAAATCACCCTCAAGAAACCTTTGGAA	1980
Db	1921	GTTTAACCTTACCGGANTATGCCAGAGATCCCCGAAAATCACCCTCAAGAAACCTTTGGAA	1980
Oy	1981	ATCCCAACCAACACTACACTACACTCAAAAGATCAGATGATCTTATTTGGCCCTTGG	2040
Db	1981	ATCCCAACCAACACTACACTACACTCAAAAGATCAGATGATCTTATTTGGCCCTTGG	2040
Oy	2041	AAACGGGATGAAACACATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGT	2100
Db	2041	AAACGGGATGAAACACATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGT	2100
Oy	2101	CCTCACCTGTGTGATACACACAGACAGACAGTATCAGTACAGACGATACAGCTGTCCACA	2160
Db	2101	CCTCACCTGTGTGATACACACAGACAGACAGTATCAGTACAGACGATACAGCTGTCCACA	2160
Oy	2161	CTGAAGAAGTGTCTCTACACAGCCCTGAATCAAAATGTTAGCTTAATAGTAAAAATCCC	2220
Db	2161	CTGAAGAAGTGTCTCTACACAGCCCTGAATCAAAATGTTAGCTTAATAGTAAAAATCCC	2220
Oy	2221	AGACTACTTCAGCCTTTAATGCTTTTATCATAAAACTGGAAGCTAGCTGAACCA	2280
Db	2221	AGACTACTTCAGCCTTTAATGCTTTTATCATAAAACTGGAAGCTAGCTGAACCA	2280
Oy	2281	TTGGAAGACATTAACCTCAAGACTCTGATTCAGAGTCGGGAAACCTTAGTCTATCTGAT	2340
Db	2281	TTGGAAGACATTAACCTCAAGACTCTGATTCAGAGTCGGGAAACCTTAGTCTATCTGAT	2340
Oy	2341	CCAAAGACGCCACACCTTAGTATGCTGCCCCAACTAAATGAGTTTAAATATCAAAATACT	2400
Db	2341	CCAAAGACGCCACACCTTAGTATGCTGCCCCAACTAAATGAGTTTAAATATCAAAATACT	2400
Oy	2401	CGTTAAAAAAGAAAAAAAAAAAAAAAAAAAAA	2429
Db	2401	CGTTAAAAAAGAAAAAAAAAAAAAAAAAAAAA	2429

Query Match	Best Local Similarity	82.9% ; Score 2013 ; DB 20 ; Length 2229 ;	Matches 2226 ; Conservative	91.6% ; Pred. No. 0 ; Mismatches 0 ; Indels 203 ; Gaps 1.
XX (MER1) MERCK & CO INC.				
PA (UYUP-) UNIV UPPSALA.				
XX Petruhkhin K, Caskey CT, Metzker M, Madellius C;				
XX WPI: 1999-540560/45.				
DR P-PSDB; AAY29953.				
XX Human and mouse polynucleotides encoding CGICE polypeptides				
PS Claim 2; Fig 2; 67pp; English.				
XX The present sequence represents the human CGICE cDNA sequence, which				
CC when mutated is responsible for Best's macular dystrophy (BMD), which				
CC polynucleotides encoding CGICE are useful for diagnosing whether a				
CC patient carries a mutation in the CGICE gene. Normal and mutated				
CC CGICE proteins are useful for identifying activators and/or inhibitors				
CC of these proteins, in order to treat BMD. The CGICE gene offers a				
CC simpler and cheaper method of diagnosing BMD without the need for the				
CC presence of the patient. The gene may also be useful to discovering				
CC the genetic cause of age-related macular dystrophy.				
XX Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 other;				
QY 1 CAGGAGTCCCAACAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 60				
DB 1 CAGGGAGTCCCAACAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 60				
QY 61 CCCACCTGACCCCAAGCCCAAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 120				
DB 61 CCCACCTGACCCCAAGCCCAAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 120				
QY 121 CAGGAGTCCCAACAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 180				
DB 121 CAGGAGTCCCAACAGCCCTAGTCGCGACGACTTCTGTGGATCATCGACCCACTCGGA 180				
QY 181 GCAGCATCTCAAGCGCTATATGGGAGTCTTATCTTCCGCTCTGCTACTACATCA 240				
DB 181 GCAGCATCTCAAGCGCTATATGGGAGTCTTATCTTCCGCTCTGCTACTACATCA 240				
QY 241 TCGCCTTATTTATAGGCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGA 300				
DB 241 TCGCCTTATTTATAGGCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGA 300				
QY 301 CTCTGTATTGCGACAGCTATATCCCATTTCTCTGCTGGCTGGCTTACG 360				
DB 301 CTCTGTATTGCGACAGCTATATCCCATTTCTCTGCTGGCTGGCTTACG 360				
QY 361 TGACGCTGTGTGACCCGCTGTGGTGAACAGTACGAACTGCGCTGGCCCGGACCGCC 420				
DB 361 TGACGCTGTGTGACCCGCTGTGGTGAACAGTACGAACTGCGCTGGCCCGGACCGCC 420				
QY 421 TCATGAGCTGTGTGCGGGCTTCTCTCAAGGCAAGGAGGCAAGGCGGGCTCTGGGCG 480				
DB 421 TCATGAGCTGTGTGCGGGCTTCTCTCAAGGCAAGGAGGCAAGGCGGGCTCTGGGCG 480				
QY 481 GCAGGCTATCCGCTACCGCAACTGGGCAACGTCATCTGCGAGGCTGACGACG 540				
DB 481 GCAGGCTATCCGCTACCGCAACTGGGCAACGTCATCTGCGAGGCTGACGACG 540				
QY 541 CAGTCTACAGCGCTTCCCGACGCGCCACGACCTGTGTGCAAGAGCGCTTATGACTCGG 600				
DB 541 CAGTCTACAGCGCTTCCCGACGCGCCACGACCTGTGTGCAAGAGCGCTTATGACTCGG 600				
QY 601 CAGAACACAAGAGTTGGAGAACTGAGCTACACACACATTTCTGGGTGCCCTGGG 660				
DB 601 CAGAACACAAGAGTTGGAGAACTGAGCTACACACACATTTCTGGGTGCCCTGGG 660				

QY 661 TGTGTTTGCCACCTGTCAATGAAAGCGTGGCTTGGAGTGCATCCGGGACCCCTATCC 720
 DB 661 TGTGTTTGCCACCTGTCAATGAAAGCGTGGCTTGGAGTGCATCCGGGACCCCTATCC 720
 QY 721 TGTCCGAGAGCTGCTGAAAGAGATGAACACCTTGGCTACTCACTGTGAGACCTGTATG 780
 DB 721 TGTCCGAGAGCTGCTGAAAGAGATGAACACCTTGGCTACTCACTGTGAGACCTGTATG 780
 QY 781 CCTACGATGATTAATATCCACTGTATATACACAGGTGTACTGTGGCGGTGTCA 840
 DB 781 CCTACGATGATTAATATCCACTGTATATACACAGGTGTACTGTGGCGGTGTCA 840
 QY 841 GCTTCTCTGACTGTCTAGTGTGGCGGAGTTTCTAACCCAGCCAGGCTTACCTG 900
 DB 841 GCTTCTCTGACTGTCTAGTGTGGCGGAGTTTCTAACCCAGCCAGGCTTACCTG 900
 QY 901 GCGATGAGCTGAGACCTGTTGTGCGGCTTCTACAGTCTTCCAGTCTTCTATGTTG 960
 DB 901 GCGATGAGCTGAGACCTGTTGTGCGGCTTCTACAGTCTTCTCCAGTCTTCTATGTTG 960
 QY 961 GCTGCTGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGATGGCCAGAGGGTTC 1020
 DB 961 GCTGCTGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGATGGCCAGAGGGTTC 1020
 QY 1021 ATGGCCAGACGTGCTTGAAGAGAGATGCAATGTCAAGAAAGAGTCTCAGCGGTAG 1080
 DB 971 ----- 970
 QY 1081 AAAGCAGCAGCGCTGTGGCGACACCTGTAAATCCAGCTACTCGGAGGCTGAGGCAAG 1140
 DB 971 ----- 970
 QY 1141 GAGATTCGCTTGAACCCGGGAGGCGGAGTGTGTGGAGAGACAGCTCATCAACCCCTT 1200
 DB 971 ----- 970
 QY 971 ----- 970
 DB 971 ----- 970
 QY 1201 TGGAGAGATGATGATGATTTTGAAGCAACTGGATTTGTGACAGCAATTTGAGAGTTC 1260
 DB 998 TGGAGAGATGATGATGATTTTGAAGCAACTGGATTTGTGACAGCAATTTGAGAGTTC 1057
 QY 1261 CCGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 DB 1058 CCGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
 QY 1321 GAATTAAGCCGAGGACAGGCGCCCTTACACAGCTGTCTCCGCCAGTTCCTGTCAGCTTC 1380
 DB 1118 GAATTAAGCCGAGGACAGGCGCCCTTACACAGCTGTCTCCGCCAGTTCCTGTCAGCTTC 1177
 QY 1381 CTTTATGGGCTCACTTCAACATCATGAGCTGTGAACAAAGAGAGATGAGATTCAGCCCAA 1440
 DB 1178 CTTTATGGGCTCACTTCAACATCATGAGCTGTGAACAAAGAGAGATGAGATTCAGCCCAA 1237
 QY 1441 TCAGAGAGACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 DB 1238 TCAGAGAGACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
 QY 1501 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
 DB 1298 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1357
 QY 1561 ATCCCTTCTCCAGAGGCGCTGCCAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 DB 1358 ATCCCTTCTCCAGAGGCGCTGCCAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
 QY 1621 CGAGGAG 1680
 DB 1418 CGAGGAG 1477
 QY 1681 GTATCAG 1740
 DB 1478 GTATCAG 1537
 QY 1741 CTTCCCTTGAAGCAATCAAGCGCGCTCAAGCTTCAAGTGTCAAGAGCATGAGACCAAA 1800

DB 1538 CTTCCCTTGAAGCAATCAAGCGCGCTCAAGCTTCAAGTGTCAAGAGCATGAGACCAAA 1597
 QY 1801 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAAA 1860
 DB 1598 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAAA 1657
 QY 1861 GAGCGATGGGCGCTTGAAGAGACCAAGAGATGATGATGATGATGATGATGATGATGATGATG 1920
 DB 1658 GAGCGATGGGCGCTTGAAGAGACCAAGAGATGATGATGATGATGATGATGATGATGATGATG 1717
 QY 1921 GTTTAACTGAGGATATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1980
 DB 1718 GTTTAACTGAGGATATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1777
 QY 1981 ATCAACAACCAATACACACTTACACTCAAGATGATGATGATGATGATGATGATGATGATGATG 2040
 DB 1778 ATCAACAACCAATACACACTTACACTCAAGATGATGATGATGATGATGATGATGATGATGATG 1837
 QY 2041 AAACGAGATTAACACACTTACACTCAAGATGATGATGATGATGATGATGATGATGATGATG 2100
 DB 1838 AAACGAGATTAACACACTTACACTCAAGATGATGATGATGATGATGATGATGATGATGATG 1897
 QY 2101 CCTACACTGTGTGATACACAGAGACACTGATGATGATGATGATGATGATGATGATGATGATG 2160
 DB 1898 CCTACACTGTGTGATACACAGAGACACTGATGATGATGATGATGATGATGATGATGATGATG 1957
 QY 2161 CTGAAGAAAGTGTCTTACACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 DB 1958 CTGAAGAAAGTGTCTTACACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2017
 QY 2221 AGACTACTGACCTTATGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
 DB 2018 AGACTACTGACCTTATGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2077
 QY 2281 TTGAAACATTTAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
 DB 2078 TTGAAACATTTAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2137
 QY 2341 CCAGAGACAGCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 DB 2138 CCAGAGACAGCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2197
 QY 2401 CCGTAAAAAAG 2429
 DB 2198 CCGTAAAAAAG 2226

RESULT 3
 ID AAV9722 standard; cDNA; 1263 BP.
 AC AAV9722;
 DT 26-Apr-1999 (first entry)
 DE Human adult retina secreted protein bk112_15 cDNA.
 XX Secreted protein; human; retina; bk112_15; ds.
 KH Homo sapiens.
 OS Homo sapiens.
 XX Key location/Qualifiers
 FH 100..885
 FT CDS /*tag- a
 ET
 XX W09856909-A2.
 XX 17-DEC-1998.
 PD 08-JUN-1998; 98WO-US11822.
 PF 05-JUN-1998; 98US-0092722.
 PR

XX	11-JUN-1997:	97US-0873218.
XX	(GEMV) GENETICS INST INC.	
PA	Agostino MJ, Fechtel K, Howes SH, Jacobs K, LaVallie ER;	
PI	McCoys JM, Racle LA, Spaulding V, Treacy M;	
PI	adult retina cDNA libraries	
XX	WPI: 1999-080899/07.	
DR	P-PSDB: AAW95345.	
XX	New polynucleotides encoding secreted human proteins - derived from	
PT	human foetal brain, adult testes, foetal kidney, adult thyroid or	
PT	adult retina cDNA libraries	
XX	Claim 12; Page 71-72; 113pp; English.	
PS		
XX	This is the nucleotide sequence of cDNA clone bk112.15, which	
CC	includes an open reading frame for a 261-amino acid polypeptide	
CC	(see AAW95345). The clone was isolated from a human adult retina	
CC	cDNA library using methods which are selective for cDNAs encoding	
CC	secreted proteins, or was identified as encoding a secreted or	
CC	transmembrane protein on the basis of computer analysis of the	
CC	amino acid sequence of the encoding protein. Database searches	
CC	indicate some sequence similarity to known sequences. The	
CC	invention provides cDNA clones (see AAW99721-33) from human adult	
CC	thyroid, adult retina, adult testis, foetal kidney and foetal brain	
CC	that encode novel secreted proteins (see AAW95344-53). Each clone is	
CC	individually available from deposit clone ATCC 96451 (see also	
CC	AAW9724-33). The isolated polynucleotides (PNS) and proteins are	
CC	predicted to have activities which would make them suitable for	
CC	treating, preventing or ameliorating medical conditions in humans	
CC	and animals, although no supporting data is given. Suggested	
CC	activities include nutritional, cytokine, cell proliferation or	
CC	differentiation, immune stimulation (e.g. as vaccines) or immune	
CC	suppressing, haematopoiesis regulating, tissue growth,	
CC	activin/inhibin, chemotactic/hemokinetic, haemostatic,	
CC	thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour	
CC	invasion suppressor, and tumour inhibition activities. The PNS are	
CC	also stated to be useful for gene therapy.	
XX		
XX	Sequence 1263 BP; 397 A; 342 C; 272 G; 252 T; 0 other:	
SQ		
	Query Match 51.2%; Score 1244.2; DB 20; Length 1263;	
	Best Local Similarity 99.8%; Pred. NO. 4.2e-280;	
	Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY	1181 GAGCAGCTTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCACTGGATGTC	1240
DB	1 GAGCAGCTTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCACTGGATGTC	60
OY	1241 GACAGGAATTTCCAGGTGTCCTGTGGCTGTGATGATGATGACCAAGACCTGCTCGG	1300
DB	61 GACAGGAATTTCCAGGTGTCCTGTGGCTGTGATGATGATGACCAAGACCTGCTCGG	120
OY	1301 ATGGAGCCGAGACATGTACTGTAAAGCCGAGACAGCCGCCCTACACAGCTGTCC	1360
DB	121 ATGGAGCCGAGACATGTACTGTAAAGCCGAGACAGCCGCCCTACACAGCTGTCC	180
OY	1361 GCCCACTTCGTCGAGCTCTTATGGCTCCACCTCAACATAGGCTGAACAAAGAG	1420
DB	181 GCCCACTTCGTCGAGCTCTTATGGCTCCACCTCAACATAGGCTGAACAAAGAG	240
OY	1421 GAGATGAGATTCCAGCCCAATAGAGAGCAAGAGATGCTCAGCTGGCATCTTGGC	1480
DB	241 GAGATGAGATTCCAGCCCAATAGAGAGCAAGAGATGCTCAGCTGGCATCTTGGC	300
OY	1481 CGCTTCCTAGGCTTCGATGCCATGATCAATCTCTCCAGGGCAACTCAAGAGCAAA	1540
DB	301 CGCTTCCTAGGCTTCGATGCCATGATCAATCTCTCCAGGGCAACTCAAGAGCAAA	360
OY	1541 CTACTGAGGCGCCAAAGAGGAATCCCTTCTCCAGAGAGGCTCCCAAAAACCAAGACA	1600
DB	361 CTACTGAGGCGCCAAAGAGGAATCCCTTCTCCAGAGAGGCTCCCAAAAACCAAGACA	420

OY	1601	GCCAACAAGAACCTTTAGGGGCCAGGAAGACAAACAAGGCGTGGAAAGCTTAAAGGCTGTGGAC	1660
Db	421	GCCAAACAGAACGTTAGGGGCGAGGAAGACAAACAAGGCGTGGAAAGCTTAAAGGCTGTGGAC	480
OY	1661	GCCCTTCAAGCTGTGGGCCACTGTATCAGAGGGCCAGGCTACTACAGTGGCCCCACAGAGGCC	1720
Db	481	GCCCTTCAAGTGTGGGCCACTGTATCAGAGGGCCAGGCTACTACAGTGGCCCCACAGAGGCC	540
OY	1721	CTCAGGCCCACTCCCATGTTCTTCTCCCTTAGAACCATTCAGGGCCGTAAAGGCTTCACAGT	1780
Db	541	CTCAGGCCCACTCCCATGTTCTTCTCCCTTAGAACCATTCAGGGCCGTAAAGGCTTCACAGT	600
OY	1781	GTCACAGGCGATAGACACCAAAAGACAAAAGCTTAAAGACTGTGAGCTTGTGGGCCCAAGAA	1840
Db	601	GTCACAGGCGATAGACACCAAAAGACAAAAGCTTAAAGACTGTGAGCTTGTGGGCCCAAGAA	660
OY	1841	AGTTTGGAAATGCTCTCAGAGAGGAGTGGGGCCCTTGATGAGCAACCAGAAAGTATCTCA	1900
Db	661	AGTTTGGAAATGCTCTCAGAGAGGAGTGGGGCCCTTGATGAGCAACCAGAAAGTATCTCA	720
OY	1901	GTGAGAGAGAAAACGTGTGGAGTTTAACTGACGGATATGCCAGAGATCCCGAAAAATCAC	1960
Db	721	GTGAGAGAGAAAACGTGTGGAGTTTAACTGACGGATATGCCAGAGATCCCGAAAAATCAC	780
OY	1961	CTCAAGAAGACCTTTGGAACTATCAACCAACCAACATACACACTACACTCAAAAGATCACATG	2020
Db	781	CTCAAGAAGACCTTTGGAACTATCAACCAACCAACATACACACTACACTCAAAAGATCACATG	840
OY	2021	GATCCTTATTTGGGCGCTTGGAAAAACAGGAGTGAAGCAATTTCTTAACCTGCTCCCTTAATGG	2080
Db	841	GATCCTTATTTGGGCGCTTGGAAAAACAGGAGTGAAGCAATTTCTTAACCTGCTCCCTTAATGG	900
OY	2081	GGATGCTTCCGCCAGCGAGGTCTCAGCTGTGTGTACACAGCAGGACACTGATCCAGTCA	2140
Db	901	GGATGCTTCCGCCAGCGAGGTCTCAGCTGTGTGTACACAGCAGGACACTGATCCAGTCA	960
OY	2141	CAGCCATACAGCTGTCCACACTGAAGAAACGTTGCTTACACAGCCTGAATCAAAATGCTTA	2200
Db	961	CAGCCATACAGCTGTCCACACTGAAGAAACGTTGCTTACACAGCCTGAATCAAAATGCTTA	1020
OY	2201	GCTTAAATAGATAAAAATCCCAACACTTCCAGCCCTTATGCTTTATTCATTAATAACT	2260
Db	1021	GCTTAAATAGATAAAAATCCCAACACTTCCAGCCCTTATGCTTTATTCATTAATAACT	1080
OY	2261	GTTAAAGCTAGAGACTGAACCATTTGAAACATTTTAACACAGCTCTGGATTCAAGTGGGA	2320
Db	1081	GTTAAAGCTAGAGACTGAACCATTTGAAACATTTTAACACAGCTCTGGATTCAAGTGGGA	1140
OY	2321	ACCCCTAGTTCATATCTGAATCCAAAGACAGCCACACCTTAGTATCTGCCAAACTAATGA	2380
Db	1141	ACCCCTAGTTCATATCTGAATCCAAAGACAGCCACACCTTAGTATCTGCCAAACTAATGA	1200
OY	2381	GTTTAAATAAATACAAATACTCGTTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT	2440
Db	1201	GTTTAAATAAATACAAATACTCGTTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT	1240
RESULT 4			
AAZ21229			
ID AAZ21229 standard; cDNA; 1916 BP.			
AAZ21229:			
22-NOV-1999 (first entry)			
Mouse CGICE cDNA sequence.			
CGICE; Best's macular dystrophy; mutation; diagnosis; detection;			
BMD; age-related macular dystrophy; ss.			
Mus sp.			

Key Location/Qualifiers
CDS 11..1666
FT /tag="a" "CGICE protein"
FT /product="CGICE protein"
XX
XX MO9943695-A1.
XX
XX 02-SEP-1999.
XX PD
XX PF 22-FEB-1999; 99MO-US03790.
XX PR 25-FEB-1998; 98US-0075941.
XX PR 18-DEC-1998; 98US-0112926.
XX
XX (MERI) MERCK & CO INC.
XX PA (UYUP-) UNIV UPPSALA.
XX PI Petrukhin K, Caskey CT, Metzker M, Wadelius C;
XX
XX WPI: 1999-540560/45.
XX DR P-PSDB: AA129955.
XX PT Human and mouse polynucleotides encoding CGICE polypeptides
XX PS Claim 2; Fig 8; 67pp; English.
XX
XX The present sequence represents the mouse CGICE cDNA sequence, which
XX when mutated is responsible for Best's macular dystrophy (BMD).
XX CC Polynucleotides encoding CGICE are useful for diagnosing whether a
XX patient carries a mutation in the CGICE gene. Normal and mutated
XX CGICE proteins are useful for identifying activators and/or inhibitors
XX of these proteins, in order to treat BMD. The CGICE gene offers a
XX simpler and cheaper method of diagnosing BMD without the need for the
XX presence of the patient. The gene may also be useful to discovering
XX the genetic cause of age-related macular dystrophy.
XX
XX Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 other:
SQ
Query Match 27.1%; Score 658.2; DB 20; Length 1916;
Best Local Similarity 67.6%; Pred. No. 2e-143;
Matches 1116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;
OY 96 TGCCTGCCATGACCATCTACTTACACAGCCAGTGGCTAATGCCGCTTGGCTCTTC 155
DB 2 TGCACAGCCATGACATCTACCTACACAAAGAGTGCATGCCGCTTGGCTCTTC 61
OY 156 TCCCGCCTGCTGCTGCTGCGGGGAGCATCTCAAGCTGCTATGCGAGTCTTA 215
DB 62 TCGTCCCTCTCTGCTGCTGCGGGAGAGCATCTCAAGCTGCTATGAGAAATTCCTT 121
OY 216 ATCTTCCTGCTGCTGCTACTACTATCATCCGCTTATTTAAGCTGCGCCTCAGGAAGA 275
DB 122 GCTTCATATTCCTCTACTACTATTCATCCGCTGACTCTACAGATGGTCTCTCGAGTGT 181
OY 276 CAACGCTGATGTTGAGAACTGACTGTATGAGACAGCTAGATCCAGCTATCC 335
DB 182 CAGCACTGTTGTTTGAAGAGCTGGCTCTGTAAGTGAAGTCAATTCAGTCAATCC 241
OY 336 ATTTCTTCTGCTGCTGCTTCTAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
DB 242 ATATCTTCTGCTGCTGCTTCTAATGTTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTT 301
OY 396 GAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
DB 302 GAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
OY 456 GAGAGCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
DB 362 GATGAGAGAGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
OY 516 CTGATCTGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
DB 422 CTGATCTGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481

OY 576 GTGCAGACAGGCTTTATGACTCCGGGAGACACAGCAGTGTGGAGAAATGACCTACCA 635
DB 482 GTGCTACAGAGTGTATGACCATGCGGAAACATGAGAGTGTGAGAAAGTGGGCGTACCA 541
OY 636 CACAAATGTTCTGGTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 695
DB 542 CACAAATGTTCTGGTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 601
OY 696 GAGAGTGAATCCGGGAGCCCTATCTGCTCCAGAGCCCTGGGAGAGGAGTGAACACCTTG 755
DB 602 GAGAGTGAATCCGGGAGCCCTATCTGCTCCAGAGCCCTGGGAGAGGAGTGAACACCTTG 661
OY 756 GCTACTGAGTGGAGCAGCTGATGCTGACAGCTGATGATGATGATGATGATGATGATGATGAT 815
DB 662 GCTACTGAGTGGAGCAGCTGATGCTGACAGCTGATGATGATGATGATGATGATGATGATGAT 721
OY 816 CAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
DB 722 CAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
OY 876 CTGAACCCAGCCAGCCCTACCCCTGATGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935
DB 782 CTGAACCCAGCCAGCCCTACCCCTGATGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
OY 936 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
DB 842 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
OY 996 GGTGAGGAGCATGGCCAGAGGGGTATGCGCCAGCAGCTGCTGATGAGAGATGAGTGA 1055
DB 877
OY 1056 CAGGAAGAGAGTCTCAGGGGTAGAAAGCAGCAGGCTGGTGGCGACACCTGTATATC 1115
DB 877
OY 1116 CCAGTACTCGGAGGCTGAGCGAGAGATGCTTGAACCCGGGAGGCGAGTGTG 1175
DB 877
OY 1176 TGCAGAGCAGCTCATCAACCCCTTGGAGAGGATGATGATGATGATGATGATGATGATGATGAT 1235
DB 879 TGCAGAGCAGCTCATCAACCCCTTGGAGAGGATGATGATGATGATGATGATGATGATGATGAT 938
OY 1236 TTGTCAGAGAGATTTGTCAGAGTCCCTGTTGCTGTGATGATGATGATGATGATGATGATGATGAT 1295
DB 939 TCATGACAGAAACCTGCAAGTCCCTGTTGCTGCTGATGATGATGATGATGATGATGATGATGAT 998
OY 1296 CTGCGATGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
DB 999 CTCCATGAGAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1058
OY 1356 CTTCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
DB 1059 CTTCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
OY 1416 AAGAGAGAGAGTTCACGCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1469
DB 1119 AAGAGAGAGTTCAGGCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1178
OY 1470 -----GATATGTTGGCGGCTTCTAGGCTGAGTGCATGATGATGATGATGATGATGATGATGAT 1523
DB 1179 ATAGAGAGCAGTATGAGCTGCTTCTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238
OY 1524 CAATCTCAAG 1583
DB 1239 AAGACTTAAAG 1292
OY 1584 CCAAAAGCAG 1643
DB 1293 GTAAGAGAGCAG 1331

Db 14096 CAAAAACCAAGGCGCCAAACAGACCTTAGGGCCGAGGAGACAAAGGCTGGAA 14155
 QY 1645 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCACTGTATCAGAGGCCAGGCTACTACAG 1704
 Db 14156 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCACTGTATCAGAGGCCAGGCTACTACAG 14215
 QY 1705 TGCCCCACAGAGGCGCCCTCAGGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 1764
 Db 14216 TGCCCCACAGAGGCGCCCTCAGGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 14275
 QY 1765 GTCAAAAGCTTACAGTGTCTACAGGCGATACACACCAAGAACAAAGCTTAAGCTGTAG 1824
 Db 14276 GTCAAAAGCTTACAGTGTCTACAGGCGATACACACCAAGAACAAAGCTTAAGCTGTAG 14335
 QY 1825 TTCTGGGGCCAAAGATTTTGAATTTGCTCTCAGAGAGCGATGGGCTTGTAGAGCA 1884
 Db 14336 TTCTGGGGCCAAAGATTTTGAATTTGCTCTCAGAGAGCGATGGGCTTGTAGAGCA 14395
 QY 1885 CCCAGAGATATCTCAAGTGTAGAGGAAAGCTGTGAGTTTAACTTACCTGACGATATGCCAGA 1944
 Db 14396 CCCAGAGATATCTCAAGTGTAGAGGAAAGCTGTGAGTTTAACTTACCTGACGATATGCCAGA 14455
 QY 1945 GATCCCCGAAATCACCCTCAAGACCTTTGAAACATCCAAACATACACTAC 2004
 Db 14456 GATCCCCGAAATCACCCTCAAGACCTTTGAAACATCCAAACATACACTAC 14515
 QY 2005 ACTCAAGATCATGATGATCTTATTTGGGCTTGGAAAGAGGATG 2051
 Db 14516 ACTCAAGATCATGATGATCTTATTTGGGCTTGGAAAGAGGATG 14562

RESULT 6
 ABA14559
 ID ABA14559 standard; DNA: 16650 BP.
 AC ABA14559;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide seq ID NO 6890.
 XX
 KW Human; noctropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antineoplastic; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antischizoid; antianemic; antitubercular; cancer;
 KW antineoplastic; hepatotoxic; cerebroprotective; antineoplastic;
 KW antiparasitic; cardiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiabetic; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 XX
 PN W0200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205151.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216847.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234937.
 PR 25-SEP-2000; 2000US-0234938.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241185.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 17-NOV-2000; 2000US-0249301.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0253978.
 (HUMAN-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-541565/60.
 Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 useful for preventing, diagnosing and/or treating nervous system
 cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 6890; 1701bp + Sequence Listing; English.
 XX
 XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias;

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 other;
 Query Match 26.4%; Score 642.2; DB 22; Length 16650;
 Best Local Similarity 99.5%; Pred. No. 2,6e-139;
 Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1405 CAGCCTGAACAAAGAGAGATGGAGTCCAGCCCAATCAGAGAGAGAGATGCTCA 1464
 DB 11895 CAGCCTGAACAAAGAGAGATGGAGTCCAGCCCAATCAGAGAGAGAGATGCTCA 11954
 OY 1465 CGGTGGCATCTTGTGGCGCTTCCAGGCTGCACTGACATCCATCTCCAGGCG 1524
 DB 11955 CGGTGGCATCTTGTGGCGCTTCCAGGCTGCACTGACATCCATCTCCAGGCG 12014
 OY 1525 AAACCTAAGAGCAAACTACTGTGTGCCCCAAGAGGGAATCCCTTCCAGAGGCTGCA 1584
 DB 12015 AAACCTAAGAGCAAACTACTGTGTGCCCCAAGAGGGAATCCCTTCCAGAGGCTGCA 12074
 OY 1585 CAAAACCAAGAGGAGGAGCAAGAACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1644
 DB 12075 CAAAACCAAGAGGAGGAGCAAGAACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12134
 OY 1645 GCTTAAGGCTGTGAGAGGCTTCAAGTGTGCCCCACATGTATAGAGGCCAGGCTACTACAG 1704
 DB 12135 GCTTAAGGCTGTGAGAGGCTTCAAGTGTGCCCCACATGTATAGAGGCCAGGCTACTACAG 12194
 OY 1705 TGCCCCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1764
 DB 12195 TGCCCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12254
 OY 1765 GTCAAGCTTACAGAGTGTGACAGGAGTACAGCAACAAAGCAAAAGCTTAAAGCTGTAG 1824
 DB 12255 GTCAAGCTTACAGAGTGTGACAGGAGTACAGCAACAAAGCAAAAGCTTAAAGCTGTAG 12314
 OY 1825 TTCTGGGGCCAAAGAAAGTTTGAATGTGCTCCACAGAGCGATGGGGCTTGATGAGCA 1884
 DB 12315 TTCTGGGGCCAAAGAAAGTTTGAATGTGCTCCACAGAGCGATGGGGCTTGATGAGCA 12374
 OY 1885 CCCAGAAGTATCTCAAGTGTGAGGAGAACTGTGAGTTTAACTGACGATATGCCAGA 1944
 DB 12375 CCCAGAAGTATCTCAAGTGTGAGGAGAACTGTGAGTTTAACTGACGATATGCCAGA 12434
 OY 1945 GATCCCGGAAATCACCCTCAAGACCTTTGGAACAATACCAACCAACATACACACTAC 2004
 DB 12435 GATCCCGGAAATCACCCTCAAGACCTTTGGAACAATACCAACCAACATACACACTAC 12494
 OY 2005 ACTCAAGATCAGATGATCTTATTGGGCTTGGAAGAGGAGT 2051
 DB 12495 ACTCAAGATCAGATGATCTTATTGGGCTTGGAAGAGGAGT 12541
 RESULT 7
 ABA14557
 ID ABA14557 standard; DNA; 18530 BP.
 XX
 AC ABA14557;
 XX
 XX 23-JAN-2002 (first entry)
 DE Human nervous system related polynucleotide SEQ ID NO 6888.
 XX Human; noctropic; neuroprotective; cytostatic; dermatological; virologic;
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;
 KW antiparkinsonian; antisticking; antianemic; antiallergic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; anticancer; anticonvulsant; antifungal;

KM anti-parasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; neprotropic; gene therapy; vaccine; ds.
 XX Homo sapiens.
 XX WO200159063-A2.
 XX 16-AUG-2001.
 XX 17-JAN-2001: 2001MO-US01334.
 XX 31-JAN-2000: 2000US-0179065.
 XX 04-FEB-2000: 2000US-0180628.
 XX 24-FEB-2000: 2000US-0184664.
 XX 02-MAR-2000: 2000US-0186350.
 XX 16-MAR-2000: 2000US-0189874.
 XX 17-MAR-2000: 2000US-0190076.
 XX 18-APR-2000: 2000US-0198123.
 XX 19-MAY-2000: 2000US-0205515.
 XX 07-JUN-2000: 2000US-0209467.
 XX 28-JUN-2000: 2000US-0214886.
 XX 30-JUN-2000: 2000US-0215135.
 XX 07-JUL-2000: 2000US-0216647.
 XX 07-JUL-2000: 2000US-0216880.
 XX 11-JUL-2000: 2000US-0217487.
 XX 11-JUL-2000: 2000US-0217496.
 XX 14-JUL-2000: 2000US-0218290.
 XX 26-JUL-2000: 2000US-0220966.
 XX 26-JUL-2000: 2000US-0220966.
 XX 14-AUG-2000: 2000US-0224518.
 XX 14-AUG-2000: 2000US-0224519.
 XX 14-AUG-2000: 2000US-0225213.
 XX 14-AUG-2000: 2000US-0225214.
 XX 14-AUG-2000: 2000US-0225266.
 XX 14-AUG-2000: 2000US-0225267.
 XX 14-AUG-2000: 2000US-0225268.
 XX 14-AUG-2000: 2000US-0225270.
 XX 14-AUG-2000: 2000US-0225447.
 XX 14-AUG-2000: 2000US-0225757.
 XX 14-AUG-2000: 2000US-0225758.
 XX 14-AUG-2000: 2000US-0225759.
 XX 18-AUG-2000: 2000US-0226279.
 XX 22-AUG-2000: 2000US-0226681.
 XX 22-AUG-2000: 2000US-0226688.
 XX 22-AUG-2000: 2000US-0227009.
 XX 23-AUG-2000: 2000US-0227009.
 XX 30-AUG-2000: 2000US-0228924.
 XX 01-SEP-2000: 2000US-0229287.
 XX 01-SEP-2000: 2000US-0229343.
 XX 01-SEP-2000: 2000US-0229344.
 XX 01-SEP-2000: 2000US-0229345.
 XX 05-SEP-2000: 2000US-0229509.
 XX 05-SEP-2000: 2000US-0229513.
 XX 06-SEP-2000: 2000US-0230437.
 XX 06-SEP-2000: 2000US-0230438.
 XX 08-SEP-2000: 2000US-0231142.
 XX 08-SEP-2000: 2000US-0231243.
 XX 08-SEP-2000: 2000US-0231244.
 XX 08-SEP-2000: 2000US-0231413.
 XX 08-SEP-2000: 2000US-0231414.
 XX 08-SEP-2000: 2000US-0232080.
 XX 08-SEP-2000: 2000US-0232081.
 XX 12-SEP-2000: 2000US-0231968.
 XX 14-SEP-2000: 2000US-0232397.
 XX 14-SEP-2000: 2000US-0232398.
 XX 14-SEP-2000: 2000US-0232399.
 XX 14-SEP-2000: 2000US-0232400.
 XX 14-SEP-2000: 2000US-0232401.
 XX 14-SEP-2000: 2000US-0233063.
 XX 14-SEP-2000: 2000US-0233064.
 XX 14-SEP-2000: 2000US-0233065.
 XX 21-SEP-2000: 2000US-0234223.
 XX 21-SEP-2000: 2000US-0234274.

PR 25-SEP-2000: 2000US-0234997.
 PR 25-SEP-2000: 2000US-0234998.
 PR 26-SEP-2000: 2000US-0235484.
 PR 27-SEP-2000: 2000US-0235834.
 PR 27-SEP-2000: 2000US-0235836.
 PR 29-SEP-2000: 2000US-0236327.
 PR 29-SEP-2000: 2000US-0236367.
 PR 29-SEP-2000: 2000US-0236368.
 PR 29-SEP-2000: 2000US-0236369.
 PR 29-SEP-2000: 2000US-0236370.
 PR 02-OCT-2000: 2000US-0236802.
 PR 02-OCT-2000: 2000US-0237037.
 PR 02-OCT-2000: 2000US-0237038.
 PR 02-OCT-2000: 2000US-0237039.
 PR 02-OCT-2000: 2000US-0237040.
 PR 02-OCT-2000: 2000US-0237041.
 PR 13-OCT-2000: 2000US-0239933.
 PR 13-OCT-2000: 2000US-0239937.
 PR 20-OCT-2000: 2000US-0240960.
 PR 20-OCT-2000: 2000US-0241785.
 PR 20-OCT-2000: 2000US-0241786.
 PR 20-OCT-2000: 2000US-0241787.
 PR 20-OCT-2000: 2000US-0241808.
 PR 20-OCT-2000: 2000US-0241809.
 PR 20-OCT-2000: 2000US-0241826.
 PR 20-OCT-2000: 2000US-0242221.
 PR 01-NOV-2000: 2000US-0244617.
 PR 01-NOV-2000: 2000US-0244674.
 PR 08-NOV-2000: 2000US-0246475.
 PR 08-NOV-2000: 2000US-0246476.
 PR 08-NOV-2000: 2000US-0246477.
 PR 08-NOV-2000: 2000US-0246478.
 PR 08-NOV-2000: 2000US-0246523.
 PR 08-NOV-2000: 2000US-0246524.
 PR 08-NOV-2000: 2000US-0246525.
 PR 08-NOV-2000: 2000US-0246526.
 PR 08-NOV-2000: 2000US-0246527.
 PR 08-NOV-2000: 2000US-0246528.
 PR 08-NOV-2000: 2000US-0246532.
 PR 08-NOV-2000: 2000US-0246609.
 PR 08-NOV-2000: 2000US-0246610.
 PR 08-NOV-2000: 2000US-0246611.
 PR 08-NOV-2000: 2000US-0246613.
 PR 17-NOV-2000: 2000US-0249207.
 PR 17-NOV-2000: 2000US-0249208.
 PR 17-NOV-2000: 2000US-0249209.
 PR 17-NOV-2000: 2000US-0249210.
 PR 17-NOV-2000: 2000US-0249211.
 PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249246.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 17-NOV-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250391.
 PR 01-DEC-2000: 2000US-0251160.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 05-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 08-DEC-2000: 2000US-0251990.
 PR 11-DEC-2000: 2000US-0254097.


```

PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239933.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

```

```

PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 05-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0254097.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases.
XX
XX Disclosure; SEQ ID NO 6889; 1701bp + Sequence Listing; English.
XX
XX PS
XX
XX The invention relates to novel genes (ABR11004-ABR21534) and proteins
XX (ABR14678-ABR18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischemia and
XX (d) wound healing; (e) neurological diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX SQ
XX Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 other;
XX
XX Query Match 26.4%; Score 640.6; DB 22; Length 18537;
XX Best Local Similarity 99.4%; Pred. No 6,4e-139;
XX Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1405 CAGCTGACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 1464
XX |||||||
XX DB 11901 CAGCTGACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 11960
XX |||||||
XX QY 1465 CGTGGCATCTTTGGCCGCTTCCAGGCGTGCAGATGATCACCATCTCCAGGGC 1524
XX |||||||
XX DB 11961 CGTGGCATCTTTGGCCGCTTCCAGGCGTGCAGATGATCACCATCTCCAGGGC 1520
XX |||||||

```

1535 AAACCTCAAGACCAACTACTGTGTGCCCCAAGAGGGAATCCCTTCTCAAGAGGGGCTGCGC 1584
12021 AAACCTCAAGACCAACTACTGTGTGCCCCAAGAGGGAATCCCTTCTCAAGAGGGGCTGCGC 12080
1585 CAAAAACCAAGAGGAGCCAAACAGACGTTAGGGGCCAGAGAACAAAGGCTGGAA 1644
12081 CAAAAACCAAGAGGAGCCAAACAGACGTTAGGGGCCAGAGAACAAAGGCTGGAA 12140
1645 GCTTAAGGCTGTGAGCGCTTCAAGTCTGCGCCACTGTATCAAGAGCCAGGCTACTACAG 1704
12141 GCTTAAGGCTGTGAGCGCTTCAAGTCTGCGCCACTGTATCAAGAGCCAGGCTACTACAG 12200
1705 TGGCCCAAGAGAGCCCGCTCAGGCCCCACCTCCATGTTCTCCCTTGAACATCAGCGCC 1764
12201 TGGCCCAAGAGAGCCCGCTCAGGCCCCACCTCCATGTTCTCCCTTGAACATCAGCGCC 12260
1765 GTCAAAAGCTTCAACAGTGTCAAGAGGATAGACACCAAAAGCTTAAAGACTGTGAG 1824
12261 GTCAAAAGCTTCAACAGTGTCAAGAGGATAGACACCAAAAGCTTAAAGACTGTGAG 12320
1825 TTCTGGGGCCCAAGAAAAGTTTAAATGCTCTCAGAGAGGAGGAGGCGCTGTATGAGCA 1884
12321 TTCTGGGGCCCAAGAAAAGTTTAAATGCTCTCAGAGAGGAGGAGGCGCTGTATGAGCA 12380
1885 CCCAGAGATATCTCAAGTGTAGAGAGAAAAGTGTGAGATTAACTGAGAGATATGCCAGA 1944
12381 CCCAGAGATATCTCAAGTGTAGAGAGAAAAGTGTGAGATTAACTGAGAGATATGCCAGA 12440
1945 GATCCCCGAAAATCAGCTCAAAAGAACCTTTGGAAACATACCAACCAACATACACATAC 2004
12441 GATCCCCGAAAATCAGCTCAAAAGAACCTTTGGAAACATACCAACCAACATACACATAC 12500
2005 ACTCAAAAGATCAGATGATCCTTATTTGGGCTTGGAAAGAGGATG 2051
12501 ACTCAAAAGATCAGATGATCCTTATTTGGGCTTGGAAAGAGGATG 12547

RESULT 9
ABL89697 standard; cDNA: 1717 BP.
ID ABL89697 standard; cDNA: 1717 BP.
AC ABL89697;
DT 24-MAY-2002 (first entry)
DE Human polynucleotide SEQ ID NO 259.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.
OS Homo sapiens.
PN WO200190304-A2.
XX 29-NOV-2001.
PD 18-MAY-2001: 2001WO-US16450.
PF 19-MAY-2000: 2000US-205515P.
PR 19-MAY-2000: 2000US-205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI: 2002-122018/16.
XX P-PDB: ABB89288.
DR Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -
XX Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
PS The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus and ulcerative
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
SO Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 9 other:
Query Match 24.2%; Score 587.8; DB 24; Length 1717;
Best Local Similarity 98.8%; Pred. No. 5e-127;
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
1445 GAGGAGAGAGAGATGCTACAGCTGCGATCATTTGGCCCTTCTGAGGCTGAGTCCAT 1504
1 GAGGAGAGAGAGATGCTACAGCTGCGATCATTTGGCCCTTCTGAGGCTGAGTCCAT 60
1505 GATCACCATCTCCCAAGGCAAACTCAAGAGCAAACTACTGTGCCCCAAGAGGGAATCC 1564
61 GATCACCATCTCCCAAGGCAAACTCAAGAGCAAACTACTGTGCCCCAAGAGGGAATCC 120
1565 CTCTCTCAGAGAGGCTGCCAAAACCAAGGAGCCAAACCAAGGAGGAGGAGGAGGAG 1624
121 CTCTCTCAGAGAGGCTGCCAAAACCAAGGAGCCAAACCAAGGAGGAGGAGGAGGAGGAG 179
1625 GAAGACAAACAGGCTGGAAGCTTAAAGCTGTGAGAGGCTTCAAGTCTGCCCACTGTAT 1684
180 GAAGACAAACAGGCTGGAAGCTTAAAGCTGTGAGAGGCTTCAAGTCTGCCCACTGTAT 239
1685 CAGAGGCGAGCTACTACAGTGTGCCCCAAGAGAGGCTTCAAGGCTTCAAGTCTGCCCACTGTAT 1744
240 CAGAGGCGAGCTACTACAGTGTGCCCCAAGAGAGGCTTCAAGGCTTCAAGTCTGCCCACTGTAT 299
1745 CCGCTAGAACATCAGGCGCGCTCAAAAGCTTCAAGTGTACAGAGGATAGACACCAAGAG 1804
300 CCGCTAGAACATCAGGCGCGCTCAAAAGCTTCAAGTGTACAGAGGATAGACACCAAGAG 359
1805 AAAAGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAAGTTTGAATGCTCTCAGAGAGC 1864
360 AAAAGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAAGTTTGAATGCTCTCAGAGAGC 419
1865 GATGGGCGCTTGTATGAGACACCAAGATATCTCAAGTGTAGAGAGGAGAAAGTGTGAGATT 1924
420 GATGGGCGCTTGTATGAGACACCAAGATATCTCAAGTGTAGAGAGGAGAAAGTGTGAGATT 479
1925 AACCTGAAGGATATGAGAGATCCCGAAATCAGCTCAAGAAAGCTTTGGAAACAATCA 1984
480 AACCTGAAGGATATGAGAGATCCCGAAATCAGCTCAAGAAAGCTTTGGAAACAATCA 539
1985 CCAACCAACATACACATACATCAAGATACATGATCTTATTTGGGCTTGGAAAC 2044
540 CCAACCAACATACACATACATCAAGATACATGATCTTATTTGGGCTTGGAAAC 599
2045 AGGAGTG 2051
600 AGGAGTG 606

RESULT 10
 ABA14556
 ID ABA14556 standard; DNA; 7108 BP.
 AC ABA14556;
 XX
 XX 23-JAN-2002 (first entry)
 DT
 DE Human nervous system related polynucleotide SEQ ID NO 6887.
 XX
 XX Human: nootropic; neuroprotective; cytoskeletal; dermatological; virocidic;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; cancer;
 KW antiparkinsonian; antisticking; antianaemic; antiarthritic; vulnery;
 KW antihemmatic; hepatocytic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antitumor;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198122.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226861.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 22-AUG-2000; 2000US-0227009.
 PR 23-AUG-2000; 2000US-0228924.
 PR 30-AUG-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 03-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236370.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 XX Disclousre; SEQ ID NO 6867; 1701pp + Sequence listing; English.
 XX
 XX The invention relates to novel genes (AB11004-ABA21534) and proteins
 CC (AB11004-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 other:
 SQ
 Query Match 19.6%; Score 476.4; DB 22; Length 7108;
 Best Local Similarity 99.4%; Pred. No. 8.2e-101;
 Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Db 6847 GCTTAAGGCTGTGAGCGCTTCAAGTCTGCCACTGATCATCAGAGCCAGGCTACTACAG 6906
 QY 1705 TGCCCCACAGAGCCCGCTCAGCCCACTCCATGTTCTTCCCTTGAACATCAGCCGC 1764
 Db 6907 TGCCCCACAGAGCCCGCTCAGCCCACTCCATGTTCTTCCCTTGAACATCAGCCGC 6966
 QY 1765 GTCAAGCTTCACTGTCATCAGAGGATATACACCAAGCAAGAGCTTAAAGACTGTGAG 1824
 Db 6967 GTCAAGCTTCACTGTCATCAGAGGATATACACCAAGCAAGAGCTTAAAGACTGTGAG 7026
 QY 1825 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTATGAGACA 1884
 Db 7027 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTATGAGACA 7086
 QY 1885 CCCAGAGTATCTCAAGTGAGG 1906
 Db 7087 CCCAGAGTATCTCAAGTGAGG 7108
 RESULT 11
 AB11094
 ID AB11094 standard; cDNA; 305 BP.
 XX
 AC AB11094;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 101.
 XX
 KW Human; nocitropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antispasmodic; antianemic; antiallergic; cancer;
 KW antihumetic; hepatotropic; cerebroprotective; antinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; neuropeptide; gene therapy; vaccine; ss.
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226861.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232403.
 PR 14-SEP-2000; 2000US-0233065.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 23-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 17-NOV-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI: 2001-541565/60.
 P-PSDB: ABBI4768.
 DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Claim 1; SEQ ID NO 101; 1701bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins
 CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia, fungal
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 305 BP; 72 A; 83 C; 86 G; 63 T; 1 other;
 Query Match 12.1%; Score 294.6; DB 22; Length 305;
 Best Local Similarity 99.7%; Pred. No. 5,6e-59;
 Matches 294; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy		1176	TGGCAGACGACGCTCATCAACCCCTTTGGAGAGAGATGATGTGATTGAGAACCACTGGG	1233
Db		1	TGGCAGACGACGCTCATCAACCCCTTTGGAGAGAGATGATGTGATTGAGAACCACTGGG	60
Qy		1236	TTGTCGCACAGGAATTTTCAGSTGCCCTGTGGCTGTGATGATGCACCAGACCTGC	1295
Db		61	TTGTCGCACAGGAATTTTCAGSTGCCCTGTGGCTGTGATGATGCACCAGACCTGC	120
Qy		1296	CTCGGATGGAGCCGCGACATGTACTGGAATAAAGCCGAGCCACAGCCCCCTTAACAAGCTG	1355
Db		121	CTCGGATGGAGCCGCGACATGTACTGGAATAAAGCCGAGCCACAGCCCCCTTAACAAGCTG	180
Qy		1356	CTTCCGCCCGCATTTCCGTCGAGCCCTCTTATGGGCTCCACCTTGACATCAGCTGAA	1415
Db		181	CTTCCGCCCGCATTTCCGTCGAGCCCTCTTATGGGCTCCACCTTGACATCAGCTGAA	240
Qy		1416	AAGAGGAGATGGATTCACGCCCAATCAGAGAGACGAGAGGATGCTCACGCTGG	1470
Db		241	AAGAGGAGATGGATTCACGCCCAATCAGAGAGACGAGAGGATGCTCACGCTGG	295
RESULT	12			
ABL10793				
ID	ABL10793	standard; cDNA; 2861 BP.		
XX	ABL10793;			
AC				
XX				
DT	26-MAR-2002	(first entry)		
XX				
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.			
XX				
KW	Drosophila; developmental biology; cell signalling; insecticide;			
XX	pharmaceutical; gene; se.			
OS	Drosophila melanogaster.			
PN	WO200171042-A2.			
XX				
PD	27-SEP-2001.			
XX				
XX	23-MAR-2001; 2001WO-US09231.			
PR	23-MAR-2000; 2000US-191637P.			
XX	11-JUL-2000; 2000US-0614150.			
PA	(PEKE) PE CORP NY.			
PI	Venter JC, Adams M, Li PMD, Myers EW;			
DR	WPI; 2001-656860/75.			
XX	P-PsDB; ABB66690.			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more			
PT	genes from Drosophila and for elucidating cell signalling and cell-cell			
PT	interactions -			
PS	Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.			
XX				
XX	The invention relates to an isolated nucleic acid detection reagent			
CC	capable of detecting 1000 or more genes from Drosophila. The invention is			
CC	useful in developmental biology and in elucidating cell signalling and			
CC	cell-cell interactions in higher eukaryotes for the development of			
CC	insecticides, therapeutics and pharmaceutical drugs. The invention			
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA			
CC	sequences (AB101840-AB116175) and the encoded proteins			
CC	(ABB57737-ABB72072).			
CC	The sequence data for this patent did not form part of the printed			
CC	specification, but was obtained in electronic format directly from WIPRO			
CC	at ftp.wipro.int/pub/published_pcr_sequences.			
XX				
Sequence	2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other:			

Query Match	9.8%;	Score 237;	DB 23;	Length 2861;
Best Local Similarity	56.6%;	Pred. No. 3,7e-45;		
Matches 510;	Conservative	0;	Mismatches 370;	Indels 21; Gaps 3;
QY	103	CCATGAGCAATCACTTACACAAAGCAAGTGG---CTAATGCCCCGCTTAGGCTCTTCC	159	
Db	413	CAATGACAAATTACGTACACAGTGAAAGTGGCCACTGTGCGGGCTTTGGCTTTTTCACA	472	
QY	160	GCGTGTCTGTCTGCGCGGGGAGCATCTACAAAGTGGTATATATGGGAGTCTTAATCT	219	
Db	473	AATGTGCTGCATATGGCCGAGGAAGATTTACAACTGGTTGGCTAGATCTTCTGGCT	532	
QY	220	TCCGCTCTCTCTACTCATCATCCGCTTATATATAGGCTGGCCCTCACGGAAGAACAC	279	
Db	533	TCCTGACCAATTTACTATGCGATACAAATGGTATATGGCTTTGGCCCAACCCGACAAA	592	
QY	280	AGCTGATGTTTGAGAAACTGACTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTT	339	
Db	593	AAGAAACCTTTGAGGCGCATTTGTCAGTACTGTATAGTTACAAAGACTCATACCCCTGT	652	
QY	340	CCTTCTGTCTGGGCTTCTAGTGAAGCGTGTGCGAACCCGCTGCTGGAACCAATACAGA	399	
Db	653	CCTTCTGTCTGGGCTTCTATGTATTCATTTGTGATGACCCGCTGGTGGAAATCAATACCT	712	
QY	400	ACCTGCGCTGGCGCCAGCCGCTCATGAGCCCTGTGCTGCGCTTCTGCAAGGCAAGAGG	459	
Db	713	CCATTCCTCCGCGCAGATGCCATCCCGGTTGTCACACTGAAATGTCATAGGCGCAGAGTG	772	
QY	460	AGCAAGCCCGCTGCTGCGGCGCACGCTCATACCGCTACGGCAACCTGGGCAAGSTGCTCA	519	
Db	773	AGCGGAGCGCATGTATGAGGAAACAAATAAAGCCATATGTGTGCTTGTGCTGACATAG	832	
QY	520	TCCCTGCGAGGGTGCAGCACCCGACTCTCAAGCCCTTCCCGAGGCCAGCACTGGTGC	579	
Db	833	TCCCTGCGAATGTTTCGCGGAGGGTGAAGAACGCTTCCCGGCTAAATATCTGGTG	892	
QY	580	AAGCAGGCTTTATGACTCCGCGAGAACACAAAGCAAGTGGGAATAGAGC-----C	630	
Db	893	AAGCGGCTCTGCTAATATGACATGAAAGAACCATATCAGAACCATGAAACAGGCTTTTC	952	
QY	631	TACACACAAACATGTTCTGGGGGCCCTGGGTGGTTGGCCAACTGTCATGAAAGCGT	690	
Db	953	CCAGACTCTGAAAGCACTGCTGCTCCATCGTTTGGGCTGCCAGTATTTAACAGGGCCA	1012	
QY	691	GCGTTGAGGTGGAATCCGGGACCTATCTGCTCCAGAGCTGCTGAACGATGAAACA	750	
Db	1013	GAAAGGAAGTGGCATTTCTGATGATTTTGGTGTGTAACCATCATCGATGATTAATA	1072	
QY	751	CCTTGGCTACTAGTGTGAGACACGCTGATAGCTTACGATGATATGCCACTGTGT	810	
Db	1073	AGTTTGTGTGATGATGTGGATGCTCTCATCAGCTACGATACCATTTAGTACTCTGGTGT	1132	
QY	811	ATACACAGGTGATGCTGGCGGTGTACAGCTTCTCCCTGACTTGTCTAGTTGGGCGCG	870	
Db	1133	ACACCCAGTGGTGAACCTGGCGGTATTCGTACTCTCTTACCTGCTGACATGGGTCAAC	1192	
QY	871	AGT-----TTTGAAACCAAGCCCAAGCCCTACCCGCGCATGAGCTGCACTGTG	921	
Db	1193	AATGAGCCGATGAGCAAGTGTGGGAAATACCAATCTGAAACAAAGGTGATATATCT	1255	
QY	922	TGCGCGCTCACGTTCTCTGCAAGTCTTCTATGTGGTGGCTGGAAGTGGGCTCT	981	
Db	1253	TTCTCTGATTTTACAAAGCTGCAAGTCTTCTTCTACATGGGTGGCTCAAGGTGGCGAGT	1312	
QY	982	C 982		
Db	1313	C 1313		

AAV6976;
27-APR-1999 (first entry)
EST clone BK112.
Expressed sequence tag; secreted protein; hematopoiesis regulator;
tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
Homo sapiens.
MO9845435-A2.
15-OCT-1998.
10-APR-1998; 98MO-US06954.
10-APR-1997; 97US-0835913.
(GEM) GENETICS INST INC.
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racie LA, Spaulding V, Treacy M;
WPI: 1999-070076/06.
New polynucleotides encoding human secreted proteins - derived from
e.g. human blood, kidney, foetal lung, placenta, testes, brain,
ovary, pituitary, retina and colon cDNA libraries
Claim 1; Page 411; 633pp; English.
This sequence represents an expressed sequence tag (EST), and is a
polynucleotide of the invention. The polynucleotides of the invention are
all secreted EST sequences isolated from a variety of human tissue
sources. The EST sequences and proteins encoded by them are predicted to
have useful biological activities which would make them suitable for
treating, preventing or ameliorating medical conditions in humans and
animals, although no supporting data is given. Suggested activities
include nutritional activity, immune stimulating or suppressing activity,
haematopoiesis regulating activity, tissue growth activity,
activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, receptor/ligand activity, anti-inflammatory
activity, cadherin/tumour invasion suppressor activity, tumour inhibition
activity. The EST sequences are also stated to be useful for gene
therapy.
Sequence 424 BP; 110 A; 124 C; 95 G; 95 T; 0 other;
Query Match 9.7%; Score 234.8; DB 20; Length 424;
Best Local Similarity 99.2%; Pred. No. 5.6e-45;
Matches 226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1814 AAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTCCTCAGAGAGCGATGGGCC 1873
DB 21 AAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTCCTCAGAGAGCGATGGGCC 80
OY 1874 TTGATGAGACCCAGAGATCTCAAGTGAGAGAGAAAGCTGTGAGTTAACTGACG 1933
DB 81 TTGATGAGACCCAGAGATCTCAAGTGAGAGAGAAAGCTGTGAGTTAACTGACG 140
OY 1934 GATATGCCAGAGATCCCGGAAATTCACCTCAAGAAACCTTGGAAACATCACCACCAAC 1993
DB 141 GATATGCCAGAGATCCCGGAAATTCACCTCAAGAAACCTTGGAAACATCACCACCAAC 200
OY 1994 ATTACACTACACTCAAGATCAATGATGATCTTATTTGGGCTTGGAAACAGGATG 2051
DB 201 ATTACACTACACTCAAGATCAATGATGATGATCTTATTTGGGCTTGGAAACAGGATG 258

ABL12609 standard; cDNA; 1608 BP.
ID ABL12609 standard; cDNA; 1608 BP.
XX ABL12609;
AC ABL12609;
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 32309.
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 32309.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
PN WO200171042-A2.
XX WO200171042-A2.
PD 27-SEP-2001.
XX 27-SEP-2001.
PF 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2001; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX 11-JUL-2000; 2000US-0614150.
PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
XX P-PSDB: ABB68506.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 32309; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL10840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1608 BP; 431 A; 336 C; 419 G; 422 T; 0 other;
Query Match 7.9%; Score 193; DB 23; Length 1608;
Best Local Similarity 52.9%; Pred. No. 5.3e-35;
Matches 468; Conservative 0; Mismatches 403; Indels 12; Gaps 2;
OY 105 ATGACATCACTTACACAGCCCAAGTGTGATGCGCGC--TTAGCTCTCTTCCGCC 161
DB 1 ATGACATCACTTACACAGCCCAAGTGTGATGCGCGC--TTAGCTCTCTTCCGCC 60
OY 162 CTGCTGTGCTGCGCGCGGAGCATCTACAGCTGTATATGCGGAGTCTTAATCTTC 221
DB 61 CTGCTGTGCTGCGCGCGGAGCATCTACAGCTGTATATGCGGAGTCTTAATCTTC 120
OY 222 CTGCTGTGCTGCGCGCGGAGCATCTATATATATATATATATATATATATATATATAT 281
DB 121 CTGCTGTGCTGCGCGCGGAGCATCTATATATATATATATATATATATATATATATAT 180
OY 282 CTGATGTTGGAAGTACTCTGTATATGCGAGCATCTACATCCAGCTATCCCATTTCC 341
DB 161 CCGTGTTCGAGGAGCATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 240
OY 342 TTGCTGTGCGGCTTACAGTACAGTGTGATGATGATGATGATGATGATGATGATGATG 401
DB 241 TTGCTGTGCGGCTTACAGTACAGTGTGATGATGATGATGATGATGATGATGATGATG 300

OY 402 CTGCGGCGCCGACCGCCCTCATGAGCGGTGTCGGGCTTCCTCCAGGCAAGGCGAG 461
 DB 301 GTTCCTTGGCCCGATCCGCGGTTATGTAGTCCCTGTCGCGCCAAAGTAA 360
 OY 462 CAGGCGCGCTGCTGCGGCGGCTCATCCGCTAGCCCAACCTGGGCAACGTGCTCATC 521
 DB 361 CATGCTGTCTGAGACGACGATTTAGATATGTGCTTGGCATTTGATGTG 420
 OY 522 CTGCGGCGGACGACCGACGATCTACAGCGCTTCCCGCGCCGACGACCTGTGCA 581
 DB 421 CTCGATGATATGCGCCAGTTATTAAGCGCTTCCACGCTACGATCATGTTAG 480
 OY 582 GCAGCGTTTATGACTCCGCGCAAGACACAGCATTTGGAGAAACGTGAC-----CTA 632
 DB 481 GTGGTTTGTCTAAACGCGCAACGAGCAATTTATGAAGCAATGATGTGAAGTTCCA 540
 OY 633 CCACAGCAACATGTTCTGGGTCGCTGGGTGTGTTGGCAACCTGTCAATGAAAGCGTGG 692
 DB 541 AAGCAGCCCAAGTATGATGATCCCATTTGTCTGGGCGCCAGATTTGTAACAAAGGCTGGA 600
 OY 693 CTGAGAGTGCATTCGGGACCTTATCTCTCTCCAGAGCTGCTGAGACGATGATACACC 752
 DB 601 AAGGAAGTGCATTTGGGATGATCTTCCCTGAAGTCCATGATGATGATGATGATGATGAT 660
 OY 753 TTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 812
 DB 661 TTCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 OY 813 ACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 872
 DB 721 ACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 OY 873 TTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932
 DB 781 TCGAT 840
 OY 933 AGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
 DB 841 AGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
 RESULT 15
 ABL10792/c
 ID ABL10792 standard; cDNA; 10760 BP.
 AC ABL10792;
 XX 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26858.
 DE Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2001; 2000US-191637P.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656660/75.
 DR P-PSDB; ABB66689.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1: SEQ ID NO 26858; 21pp + Sequence Listing; English.
 XX
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB161175) and the encoded proteins
 CC (AB161737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 10760 BP; 3232 A; 2079 C; 2173 G; 3276 T; 0 other;
 SQ

Query Match 7.7%; Score 188.2; DB 23; Length 10760;
 Best Local Similarity 56.4%; Pred. No. 1,5e-33;
 Matches 402; Conservative 0; Mismatches 293; Indels 18; Gaps 2;

OY 288 TTGAGAACTGACTGCTGATTCGACAGCTACATCCAGCTGATCCCATTTCTTGTG 347
 DB 506 TTGAGGCCATTGTTTACGATCTGTATGATGAGAACTGATACCTGCTTCCTG 5009
 OY 348 CTGCGCTTCTACGAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 407
 DB 5008 CTGCGCTTCTACGAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 4949
 OY 408 TTGCGCCACCGCCGCTACGAGCTGCTGCTGAGAACCAATGACCAACCGCG 467
 DB 4948 TTGCGCCACCGCCGCTACGAGCTGCTGCTGAGAACCAATGACCAACCGCG 4889
 OY 468 CGGCTGCTGCGCGCAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 527
 DB 4888 CGGCTGCTGCGCGCAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 4829
 OY 528 AGCTGACAGCCGAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 587
 DB 4828 AGCTGACAGCCGAGCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 4769
 OY 588 TTATGACTCGCGGACAGCAACAGCTGAGAACTGAGC-----CTACACAC 638
 DB 4768 TTATGACTCGCGGACAGCAACAGCTGAGAACTGAGC-----CTACACAC 4709
 OY 639 AACATGTTGCGGTCCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 698
 DB 4708 AACATGTTGCGGTCCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 4649
 OY 699 GGTGCAATCGGAGACCTATCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 758
 DB 4648 GGTGCAATCGGAGACCTATCTGCTGCTGACCCGCTGCTGAGAACCAATGACCAACCGCG 4589
 OY 759 ACTGATGAGACACCTGATGCTGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 818
 DB 4588 ACTGATGAGACACCTGATGCTGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 4529
 OY 819 GTGCGTACGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 DB 4528 GTGCGTACGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4469
 OY 874 ----TTCTGAACCCAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
 DB 4468 GTGCGTACGCTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4409
 OY 930 TTACAGCTTCTGAGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 982
 DB 4408 TTACAGCTTCTGAGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4356

Mon Jun 9 10:49:30 2003

Search completed: June 8, 2003, 11:23:51
Job time : 492.416 secs

us-09-622-964-4.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 8, 2003, 09:30:48 ; Search time 6124.06 Seconds
(without alignments)
11543.117 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429

Sequence: 1 caggagagtcacacacaccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hg_hum:*

31: em_hg_inv:*

32: em_hg_inv:*

33: em_hg_mus:*

34: em_hg_pln:*

35: em_hg_rod:*

36: em_hg_mam:*

37: em_hg_vrt:*

38: em_sy:*

39: em_hggo_hum:*

40: em_hggo_mus:*

41: em_hggo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2418.4	99.6	2420	9 AF057170	AF057170 Homo sapi
2	1995.4	82.1	2210	9 AF057169	AF057169 Homo sapi
3	1953.2	80.4	2171	9 AF073501	AF073501 Homo sapi
4	648.4	26.7	1289	4 AY064707	AY064707 Sus scro
5	642.2	26.4	142092	2 AF139813	AF139813 Homo sapi
6	642.2	26.4	163915	2 AF139813	AF139813 Homo sapi
7	642.2	26.4	166867	9 AC087451	AC087451 Homo sapi
8	642.2	26.4	196080	9 AC004228	AC004228 Homo sapi
9	640.6	26.4	112309	9 AC003025	AC003025 Homo sapi
10	640.6	26.4	133683	2 AC084857	AC084857 Homo sapi
11	552.8	22.8	180169	9 AC051664	AC051664 Homo sapi
12	431	17.7	1908	9 AF440756	AF440756 Homo sapi
13	415	17.1	1956	10 BC031186	BC031186 Mus muscu
14	415	17.1	1957	10 BC019528	BC019528 Mus muscu
15	389.2	16.0	1861	9 AK000139	AK000139 Homo sapi
16	389.2	14.8	492	9 HSYMD2P11	AF073500 Homo sapi
17	358.4	14.0	106648	9 AC004588	AC004588 Homo sapi
18	339.4	14.0	2045	9 AF440757	AF440757 Homo sapi
19	338.2	13.9	1506	9 AF440758	AF440758 Homo sapi
20	338.2	13.9	2500	9 AK096459	AK096459 Homo sapi
21	338.2	13.9	2500	9 HSYMD2P04	AF073493 Homo sapi
22	240.8	9.9	413	9 HSYMD2P04	AF052095 Homo sapi
23	240.8	9.9	1666	9 AF052095	AF052095 Homo sapi
24	237	9.8	2874	3 AF218817	AF218817 Drosophi
25	237	9.8	5610	3 AY061546	AY061546 Drosophi
26	232.8	9.6	295	9 HSYMD2P07	AF073496 Homo sapi
27	191.6	7.9	246	10 AF057171	AF057171 Mus muscu
28	189	7.8	272	9 HSYMD2P02	AF073491 Homo sapi
29	189	7.8	160169	2 AC051664	AC051664 Homo sapi
30	188.2	7.7	32022	2 AC019521	AC019521 Drosophi
31	188.2	7.7	170675	3 AC008139	AC008139 Drosophi
32	188.2	7.7	192055	3 AC009183	AC009183 Drosophi
33	188.2	7.7	221888	3 AE003686	AE003686 Drosophi
34	175	7.2	438	11 G05922	G05922 human SMS W
35	175	7.2	1119	9 H05922	H05922 Human ferr
36	175	7.2	1198	6 AX409527	AX409527 Sequence
37	175	7.2	1198	9 H05922	L20941 Human ferr
38	161.4	6.6	239	9 HSYMD2P05	AF073494 Homo sapi
39	155	6.4	125804	2 AC019853	AF073498 Homo sapi
40	150.4	6.2	163514	3 AC091227	AC019853 Drosophi
41	150.4	6.2	274351	3 AE003531	AC091227 Drosophi
42	150.4	6.2	274351	3 AE003531	AE003531 Drosophi
43	135.2	5.6	147750	9 AC018761	AC018761 Homo sapi
44	130.4	5.4	90487	9 AL592166	AL592166 Human DNA
45	130.4	5.4	170141	2 AL611928	AL611928 Homo sapi

ALIGNMENTS

RESULT 1
AF057170 2420 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
complete cds.
VERSION AF057170.1 GI:3335160
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2420)
Petrushkin, K., Koistinen, M.J., Bakali, B., Li, W., Xie, G., Marknell, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,

Pred. No. 1s the number of results predicted by chance to have a

Bergen, A.A., McGarty-Dugan, V., Figueroa, D., Austin, C.P.,
Metzker, M.L., Caskey, C.T. and Wadellus, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
JOURNAL 98324772
PUBMED 9663395
REFERENCE 2 (bases 1 to 2420)
AUTHORS Petrakian, K.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
source Location/Qualifiers
1..2420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2420
/gene="VMD2"
105..1412
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform; contains exon 7"
/product="bestrophin"
/protein_id="AAC64344.1"
/db_xref="GI:3335161"
/translation="MTTYSQYANRIGSFRLICMRSIYKLYEFLLECY
IRFIYRLALTEQQLMEKTLICDSYDILIPISYLVGYATLVYTRMNOVENPM
PDRMLVSGFVKDEQKQRLRLRLILRYANLNLNLSVSTAYVKKRPSNOHLOA
GPMTPAHKOLEKISLPHNMFVPMWVPMANLSKAWLGGHIDPILQSLINMPLR
TCGHILAYDMISIPLYTVAVVSFLTCLVGRQFNPAKAYPEHLLDVVE
TRLOFFYVMKIVKLSRALGMHGRGHGOQLLETMRMCOERKYSRVESLDMWT
VLPATREARAGESLPERGRIRIMOSSSTPLEMMMLLPTGLSTGICRCPLMNR
CTRTCLMSRCTGTSIPSHSPPIOLLPSSVPEPLMAPETSA"
BASE COUNT 605 a 695 c 612 g 508 t
ORIGIN
Query Match 99.6%; Score 2418.4; DB 9; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CAGGAGTCCACACAGCTAGTCCGACAGCTTCTGTGGATCATCCGACCTGGAA 60
DB 1 CAGGAGTCCACACAGCTAGTCCGACAGCTTCTGTGGATCATCCGACCTGGAA 60
QY 61 CCCACCTGACCAAGCCCACTGCTGAGCCACCTGCTGACATGACCACTTACA 120
DB 61 CCCACCTGACCAAGCCCACTGCTGAGCCACCTGCTGACATGACCACTTACA 120
QY 121 CAAAGCAAGTGGCTAATGCCCGCTTAAAGCTCTCTCCGCTGCTGCTGCTGCTG 180
DB 121 CAAAGCAAGTGGCTAATGCCCGCTTAAAGCTCTCTCCGCTGCTGCTGCTGCTG 180
QY 181 GCAGCATCTACAAAGCTGCTATATGGGAGTCTTATCTTCCGCTCTCTACTACATCA 240
DB 181 GCAGCATCTACAAAGCTGCTATATGGGAGTCTTATCTTCCGCTCTCTACTACATCA 240
QY 241 TCCGCTTATTTATATAGGCTGCGCCCTACAGGAAGACAGAGCTGATGTTGAGAAACTCA 300
DB 241 TCCGCTTATTTATATAGGCTGCGCCCTACAGGAAGACAGAGCTGATGTTGAGAAACTCA 300
QY 301 CTCTGATTTGCGACAGCTATCATCACTCATCCCAATTCCTTCTGCTGCTGCTGCTGCTG 360
DB 301 CTCTGATTTGCGACAGCTATCATCACTCATCCCAATTCCTTCTGCTGCTGCTGCTGCTG 360
QY 361 TGAGCGTGGTGTGACCCGCTGCTGGAACAGTACAGAACTGCGCTGCGCCGACCGCC 420
DB 361 TGAGCGTGGTGTGACCCGCTGCTGGAACAGTACAGAACTGCGCTGCGCCGACCGCC 420
QY 421 TCATGAGCTGTGTCGAGCTTCTGCAAGAGCAAGCAAGAGCGCTGCTGCTGCTGCTG 480
DB 421 TCATGAGCTGTGTCGAGCTTCTGCAAGAGCAAGCAAGAGCGCTGCTGCTGCTGCTGCTG 480

QY 481 GCAGCTCATTCGCTACAGCAACCTGGGCAAGCTGCTCATCTCTGCGCAGCTGACACCG 540
DB 481 GCAGCTCATTCGCTACAGCAACCTGGGCAAGCTGCTCATCTCTGCGCAGCTGACACCG 540
QY 541 CAGCTACAAAGGCTTCCCGACGCGCCACACCTGCTGACAGCAGCTTATGATCTCCGG 600
DB 541 CAGCTACAAAGGCTTCCCGACGCGCCACACCTGCTGACAGCAGCTTATGATCTCCGG 600
QY 601 CAGAACACAAAGCTTGGAGAACTGAGCCTACCAACAAATGTTCTGGTGCCCTGGG 660
DB 601 CAGAACACAAAGCTTGGAGAACTGAGCCTACCAACAAATGTTCTGGTGCCCTGGG 660
QY 601 CAGAACACAAAGCTTGGAGAACTGAGCCTACCAACAAATGTTCTGGTGCCCTGGG 660
QY 661 TGTGCTTGGCCCAACCTGCTCAATGAGGCTGCTGAGGTGCAATCCGGACCTATCC 720
DB 661 TGTGCTTGGCCCAACCTGCTCAATGAGGCTGCTGAGGTGCAATCCGGACCTATCC 720
QY 721 TGTCTCAGAGCTCTGTAAGAGATGAACACTTGTGCTACTCATGCTGAGACCTGATAG 780
DB 721 TGTCTCAGAGCTCTGTAAGAGATGAACACTTGTGCTACTCATGCTGAGACCTGATAG 780
QY 781 CCTACGACTGATTAATGATCCCACTGGTATACAGAGTGTGACTGTGGCGGTATCA 840
DB 781 CCTACGACTGATTAATGATCCCACTGGTATACAGAGTGTGACTGTGGCGGTATCA 840
QY 841 GCTTCTCTGACTTGTCTAGTTGGCGGCACTTCTGACACCCAGCCCAAGGCTTACCCTG 900
DB 841 GCTTCTCTGACTTGTCTAGTTGGCGGCACTTCTGACACCCAGCCCAAGGCTTACCCTG 900
QY 901 GCGATGAGCTGAGCTCTGTTGCGCCGCTTACAGTCTCTGAGTCTTCTATATGTTG 960
DB 901 GCGATGAGCTGAGCTCTGTTGCGCCGCTTACAGTCTCTGAGTCTTCTATATGTTG 960
QY 961 GCTGGGTGAAGTGGGCGCTCTGACAGGCGCTGCTGCTGCTGAGCATGCGCAGAGGGTCT 1020
DB 961 GCTGGGTGAAGTGGGCGCTCTGACAGGCGCTGCTGCTGCTGAGCATGCGCAGAGGGTCT 1020
QY 1021 ATGGCCAGCAGCTGCTTGAAGAGAGATGCACTGTCAGAAAGAGGCTCAGCGGTAG 1080
DB 1021 ATGGCCAGCAGCTGCTTGAAGAGAGATGCACTGTCAGAAAGAGGCTCAGCGGTAG 1080
QY 1081 AAGGACAGCAGGCTGCTGAGGAGCAGCTGTAATCCCACTCTGCGGAGCTGAGGAG 1140
DB 1081 AAGGACAGCAGGCTGCTGAGGAGCAGCTGTAATCCCACTCTGCGGAGCTGAGGAG 1140
QY 1141 GAGAACTGCTGAACCCCGGAGGAGGAGGCTGCTGCTGAGAGCAAGCTATCAACCCCTT 1200
DB 1141 GAGAACTGCTGAACCCCGGAGGAGGAGGCTGCTGCTGAGAGCAAGCTATCAACCCCTT 1200
QY 1201 TGGAGAGATGATGATGATTTTGAAGCAACCTGATTTGACAGAGAAATTTGAGGTGTC 1260
DB 1201 TGGAGAGATGATGATGATTTTGAAGCAACCTGATTTGACAGAGAAATTTGAGGTGTC 1260
QY 1261 TGGAGAGATGATGATGATTTTGAAGCAACCTGATTTGACAGAGAAATTTGAGGTGTC 1320
DB 1261 TGGAGAGATGATGATGATTTTGAAGCAACCTGATTTGACAGAGAAATTTGAGGTGTC 1320
QY 1321 GAATTAAGCCGAGCAGCAAGCCCTTACAGAGCTTCCGCGAGTCCGCTGAGAGCTTC 1380
DB 1321 GAATTAAGCCGAGCAGCAAGCCCTTACAGAGCTTCCGCGAGTCCGCTGAGAGCTTC 1380
QY 1381 CTTTATGAGCTTCAACCTTCAACATCAGCCTGTAACAAAGAGAGATGAGTTCAGGCCAA 1440
DB 1381 CTTTATGAGCTTCAACCTTCAACATCAGCCTGTAACAAAGAGAGATGAGTTCAGGCCAA 1440
QY 1441 TCAGAGAGAGAGAGAGATGCTCACGCTGAGCATATTTGGCGCTTCTAGAGCTGCACTC 1500
DB 1441 TCAGAGAGAGAGAGAGATGCTCACGCTGAGCATATTTGGCGCTTCTAGAGCTGCACTC 1500
QY 1501 CCATGATACCACTCTCCAGGCGCAACTCAAGGACCAACTACTGTTGGCCCAAGAGGA 1560
DB 1501 CCATGATACCACTCTCCAGGCGCAACTCAAGGACCAACTACTGTTGGCCCAAGAGGA 1560

Db 361 TACGCTGCTGATGACCCGCTGCTGTAACAGTACAGAGAACTGCGCTGCGCCGACGCGC 420
Qy 421 TCATGAGCCTGCTGCTGCGCTTCTGTAAGCAAGAGACAGCAAGGCGCGCTGCTGCGC 480
Db 421 TCATGAGCCTGCTGCTGCGCTTCTGTAAGCAAGAGACAGCAAGGCGCGCTGCTGCGC 480
Qy 481 GCACGCTCATCCGCTACGCGCAACCTGCGGCAAGGCTGCTCATCCGCGCAGCTGACAGCG 540
Db 481 GCACGCTCATCCGCTACGCGCAACCTGCGGCAAGGCTGCTCATCCGCGCAGCTGACAGCG 540
Qy 541 CAGTACAGAGCCTTCCCGAGCGCCAGACCTGCTGCAAGCAGGCTTTATGACTCCG 600
Db 541 CAGTACAGAGCCTTCCCGAGCGCCAGACCTGCTGCAAGCAGGCTTTATGACTCCG 600
Qy 601 CAGACACAGCAGCTGAGAGAACTGAGCTTACCAACAACTGCTGCTGCTGCTGCTGCTG 660
Db 601 CAGACACAGCAGCTGAGAGAACTGAGCTTACCAACAACTGCTGCTGCTGCTGCTGCTG 660
Qy 661 TGTGCTTGCACACCTGCTCAATGAAAGGCTGCTGAGAGTCAATCCGCGACCTATCC 720
Db 661 TGTGCTTGCACACCTGCTCAATGAAAGGCTGCTGAGAGTCAATCCGCGACCTATCC 720
Qy 721 TGTGCTCAGAGCCTGCTGAGAGAGATGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 TGTGCTCAGAGCCTGCTGAGAGAGATGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 781 CCTACAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 CCTACAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 841 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 ATGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 971 ----- 970
Qy 1081 AAGCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 971 ----- 970
Qy 1141 GAGAACTGCTGTAACCGGAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 971 ----- 970
Qy 1201 TGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 998 TGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1057
Qy 1261 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1058 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117
Qy 1321 GAATAAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1380
Db 1118 GAATAAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1177
Qy 1381 CTTTATGAGGCTCACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1440
Db 1178 CTTTATGAGGCTCACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1237
Qy 1441 TCAGAGAGAGAGAGAGATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 TCAGAGAGAGAGAGAGATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

Db 1238 TCAGAGAGAGAGAGAGATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
Qy 1501 CCATGATCACCCTCTCCAGGCGCAACTCAAGAGCAAACTATGCTGGCCAGAGAGGA 1560
Db 1298 CCATGATCACCCTCTCCAGGCGCAAACTCAAGAGCAAACTATGCTGGCCAGAGAGGA 1357
Qy 1561 ATCCCTTCTCCAGAGGCGCTGCGCAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1358 ATCCCTTCTCCAGAGGCGCTGCGCAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
Qy 1621 CCAG 1680
Db 1418 CCAG 1477
Qy 1681 GATACAGAGGCGCCTACTACAGTGGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1478 GATACAGAGGCGCCTACTACAGTGGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
Qy 1741 CTTCCCTTGAAGAGATGAG 1800
Db 1538 CTTCCCTTGAAGAGATGAG 1597
Qy 1801 AGACAAAAGCTTAAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1598 AGACAAAAGCTTAAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657
Qy 1861 GAGCAGTGGGCGCTGATGAG 1920
Db 1658 GAGCAGTGGGCGCTGATGAG 1717
Qy 1921 GTTAACTGAGAGATGATGAG 1980
Db 1718 GTTAACTGAGAGATGATGAG 1777
Qy 1981 ATACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 1778 ATACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1837
Qy 2041 AAACAGAGATGAG 2100
Db 1838 AAACAGAGATGAG 1897
Qy 2101 CCTACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 1898 CCTACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1957
Qy 2161 CTGAAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 1958 CTGAAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2017
Qy 2221 AGACTACTGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Db 2018 AGACTACTGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2077
Qy 2281 TTGGAACATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db 2078 TTGGAACATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2137
Qy 2341 CCAAG 2400
Db 2138 CCAAG 2197
Qy 2401 CGTTAAAAAAA 2413
Db 2198 CGTTAAAAAAA 2210

RESULT 3
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA.
DEFINITION complete cds.
ACCESSION AF073501

VERSION AF073501.1 GI:3511241
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1. (bases 1 to 2171)
 AUTHORS Stohr, H., Marguardt, A., Rivera, A., Cooper, P.R., Nowak, N.J., Shows, T.B., Gerhard, D.S. and Weber, B.H.
 TITLE A gene map of the Best's vitelliform macular dystrophy region in chromosome 11q12-q13.1
 JOURNAL Genome Res. 8 (1), 48-56 (1998)
 MEDLINE 98112782
 PUBMED 9445487
 REFERENCE 2. (bases 1 to 2171)
 AUTHORS Marguardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A. and Weber, B.H.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUN-1998) Human Genetics, University, Bionzentrum, Am Hubland, Wurzburg 97074, Germany
 FEATURES
 source
 1. 2171
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q12-q13.1"
 /tissue_type="retinal pigment epithelium"
 1. 2171
 /gene="VMD2"
 89. 1846
 /gene="VMD2"
 /note="Best disease"
 /codon_start=1
 /product="vitelliform macular dystrophy protein"
 /protein_id="AAC33766.1"
 /db_xref="GI:3511242"
 /translation="MTITVSVANARLGSFRLLCWNGSIYKLLYGFLLIFLICY IIRFIYRLALTEEOQLMEKFLKLYCDSTYOLIPISFVLFYTLVTVRWNOYENLPW PDRMLVSGFEVGEQKDEGRLLRLIRYANGNVLILRSVSTAVYKRPSQHLVQA GMPTRAEHKLKLSLPHMFWPMWVFNLSKAWLGRIDPILLOSLIEMNTLR TQCGHLKATMDKLSIPVYQVTVAVSFPLICVGRFLNPAKAVPGHELDLVYVE FLQGFYVGMKLVAVQOLINPREDDEDEETWNIYDRNLOVSLIANDENHODLPHEP DMYWNKPEQPPYTAASQFRASFEWSTFNISLNEKEEPOPNODEDNAMKILAV FLGLQSHDHPRPANRSRTKLMPKRESLHEGLPKNHKAKOVNRQEDNKMKLKAV DAKSAPLVORPGYYSAPQPLSPFPLPSPSKLSVYGLDTPKXSLKTYVSSG AKKSPFLSSDGLMEHPESQVRRKRYEFLNLPMEIPEHNLKPELROSPFNHTT LKDHMPYALLENRDEAHS"
 BASE COUNT 540 a 638 c 523 g 470 t
 ORIGIN
 Query Match 80.4%; Score 1953.2; DB 9; Length 2171;
 Best Local Similarity 91.3%; Pred. No. 0;
 Matches 2168; Conservative 0; Mismatches 3; Indels 203; Gaps 1;

317 CTACATCCAGCTCATCCCATTTCTTGTGCTGAGCTTACGTGACGCTGCTGAC 376
 301 CTACATCCAGCTCATCCCATTTCTTGTGCTGAGCTTACGTGACGCTGCTGAC 360
 377 CCGCTGCTGAGAACACATACAGACAGACCTGCGGCGGACCGCCATACAGCTGCTG 436
 361 CCGCTGCTGAGAACACATACAGACAGACCTGCGGCGGACCGCCATACAGCTGCTG 420
 437 GGGCTGCTGAGAACACATACAGACAGACCTGCGGCGGACCGCCATACAGCTGCTG 496
 421 GGGCTGCTGAGAACACATACAGACAGACCTGCGGCGGACCGCCATACAGCTGCTG 480
 497 CCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACACCGCATAGTCTAAGCGCT 556
 481 CCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACACCGCATAGTCTAAGCGCT 540
 557 CCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACACCGCATAGTCTAAGCGCT 616
 541 CCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACACCGCATAGTCTAAGCGCT 600
 617 GGAAACCTGAGCTACACACACACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 676
 601 GGAAACCTGAGCTACACACACACATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 677 GTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
 661 GTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 737 GAACGATGATACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
 721 GAACGATGATACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 797 TATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
 781 TATCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 857 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
 841 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 917 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
 901 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
 977 CCTCTCAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
 955 ----- 954
 1037 TGAGACGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
 955 ----- 954
 1097 GTGGCGACACTGTATCCAGCTATCTGCGGAGGCTGAGGAGAGAAATGCTTGAACC 1156
 955 ----- 954
 1157 CCGGAGCGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
 955 ----- 954
 1217 GATTTGAGACCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276
 998 GATTTGAGACCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
 1058 GAGATGACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
 1277 GAGATGACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1336
 1337 CAGCCCTTACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396
 1118 CAGCCCTTACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177

1397 TTCAACATCAGCCGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1456
 1178 TTCAACATCAGCCGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1237
 1457 GATGCTCAGCGTGCATGATGAGCGGCTCTAGGCTGAGTCCATGATCACCATT 1516
 1238 GATGCTCAGCGTGCATGATGAGCGGCTCTAGGCTGAGTCCATGATCACCATT 1287
 1517 CCCAGGCAAACTCAAGAGCAAACTAGTGTGGCCCAAGAGGAAATCCCTTCTCCAG 1576
 1298 CCCAGGCAAACTCAAGAGCAAACTAGTGTGGCCCAAGAGGAAATCCCTTCTCCAG 1357
 1577 GGCCTGCCCAAAACCAAGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1636
 1358 GGCCTGCCCAAAACCAAGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1417
 1637 GGCCTGCCCAAAACCAAGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1696
 1418 GGCCTGCCCAAAACCAAGGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1477
 1697 TACTACAGTCCCGCCACAGAGCCCGCTCAGCCCGCTCAGCCCGCTCAGCCCGCT 1756
 1478 TACTACAGTCCCGCCACAGAGCCCGCTCAGCCCGCTCAGCCCGCTCAGCCCGCT 1537
 1757 TCAAGCCCGCTCAGAGCTTCAAGTGTGACAGGACATGACCAAGCAAGCAAGCAAG 1816
 1538 TCAAGCCCGCTCAGAGCTTCAAGTGTGACAGGACATGACCAAGCAAGCAAGCAAG 1597
 1817 ACTGAGTGTCTGGGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTG 1876
 1598 ACTGAGTGTCTGGGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTG 1657
 1877 ATGAGAGCCCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1936
 1658 ATGAGAGCCCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1717
 1937 ATGAGAGCCCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1996
 1718 ATGAGAGCCCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1777
 1997 CACACTACACTCAAAAGATACATGATGATCTTATGGGCGCTTGGAAAACAGGAGTGA 2056
 1778 CACACTACACTCAAAAGATACATGATGATCTTATGGGCGCTTGGAAAACAGGAGTGA 1837
 2057 CATCTCAACTGCTTCTTCAATGAGAGTCTTCCAGAGCGAGTCTCAGCTGTGTATC 2116
 1838 CATCTCAACTGCTTCTTCAATGAGAGTCTTCCAGAGCGAGTCTCAGCTGTGTATC 1897
 2117 ACCAGCAGAGACAGTATCAGATCAGAGCAGTATGAGTCTCAGACTGAAAGAGTGTCT 2176
 1898 ACCAGCAGAGACAGTATCAGATCAGAGCAGTATGAGTCTCAGACTGAAAGAGTGTCT 1957
 2177 ACCAGCAGAGACAGTATCAGATCAGAGCAGTATGAGTCTCAGACTGAAAGAGTGTCT 2236
 1958 ACCAGCAGAGACAGTATCAGATCAGAGCAGTATGAGTCTCAGACTGAAAGAGTGTCT 2017
 2237 TATGCTCTTTATTCATATAAACTGTGAAGCTAGACTGAAACATTTAACT 2296
 2018 TATGCTCTTTATTCATATAAACTGTGAAGCTAGACTGAAACATTTAACT 2077
 2297 CAGACTGTGATTCAGAGTCCGGGAACTTATGTTATCTGAAATCAAGAGCAGCAGC 2356
 2078 CAGACTGTGATTCAGAGTCCGGGAACTTATGTTATCTGAAATCAAGAGCAGCAGC 2137
 2357 TTAGATATCTGCCCAAACTAATGAGTTTAAATAA 2390
 2138 TTAGATATCTGCCCAAACTAATGAGTTTAAATAA 2171

RESULT 4
 LOCUS AY064707 1289 bp mRNA linear MM 19-AUG-2002
 DEFINITION Sus scrofa domestica mRNA, partial cds.

ACCESSION AY064707 GI:18476495
 VERSION AY064707.1
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabbs, J.W.
 and Marmorstein, A.D.
 Bestrophin interacts physically and functionally with protein
 phosphatase 2A
 J. Biol. Chem. 277 (34), 30591-30597 (2002)
 12058047
 2 (bases 1 to 1289)
 Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabbs, J.W.
 and Marmorstein, A.D.
 Direct Submission
 Submitted (30-NOV-2001) Cole Eye Institute, 131, Cleveland Clinic
 Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
 FEATURES
 source
 CDS
 1..1289
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 <1..1289
 /product="bestrophin"
 /protein_id="AA040882.1"
 /db_xref="GI:18476496
 /translation="POHLVKGAFMTPSHKLHKLHSLPHNSFMRPMTYFANLSIKAWI
 GRRIDPVLQSLIDEMNTLFTFCGLHAYWMTISVPLTVTVVAVYSFLACVGR
 OFNLAKVYVGHMDLVVPLFTFLFFAYAMLYKAEQILNFGDDDDFFETNMLVDR
 SIQVSLISAVDEHNDLHPRMDMYNDEPHPPYAAQSRSEFSGSTNLSLGR
 DVEFQPEEEKRNHGLILGHILGSLDSDOPRNTKTLNPKKGHEHSHKPKRLV
 ARLDSSDDSDSKAMREGEKSLAGRCGYSAPOPTGLHPRMYPRESAPLGLRV
 SGIDBAANDQSLQPAIPSTIKSKFELPPSAEASAPPLDGSVHRRTVEFNLAADSEAP
 EHLKEPNLEPMPGIIHALHNDHPYALENRDEANS"
 BASE COUNT 293 a 390 c 360 g 246 t
 ORIGIN
 Query Match 26.7%; Score 648.4; DB 4; Length 1289;
 Best Local Similarity 70.3%; Pred. No. 2e-143;
 Matches 1058; Conservative 0; Mismatches 221; Indels 227; Gaps 5;
 565 CCCAGACCTGTGTCAGAGAGGCTTATGACTCCGAGCAACAAAGCAGTGGAGAAAC 624
 4 CCCAGACCTGTGTCAGAGAGGCTTATGACTCCGAGCAACAAAGCAGTGGAGAAAC 63
 625 TGAAGCTTACCAACAAAGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 684
 64 TGAAGCTTACCAACAAAGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 123
 685 AGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 744
 124 AGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 183
 745 TGAACACCTTGCCTACTAGTGTGACACCTGTATGCTTACAGATGATAGATCCAC 804
 184 TGAACACCTTGCCTACTAGTGTGACACCTGTATGCTTACAGATGATAGATCCAC 243
 805 TGGTGTATACAGAGTGTGACAGTGTGGGAGTACAGTCTTCTCTGACTTGTCTAGT 864
 244 TGGTGTATACAGAGTGTGACAGTGTGGGAGTACAGTCTTCTCTGACTTGTCTAGT 303
 865 GGGCGAGTTTCTGAACCCAGCAAGCCTACCTGGCCATGAGCTGAGCTGCTGTGTC 924
 304 GGAGCAGTTTCTGAACCCAGCAAGCCTACCTGGCCATGAGCTGAGCTGCTGTGTC 363
 925 CCGTCTTACAGTCTTCTGAGTCTTCTTCTTATGTTGGCTGGCTGAGAGTGGGCTTCA 984
 364 CCGTCTTACAGTCTTCTGAGTCTTCTTCTTATGCTTCTTCTTCTTCTTCTTCTTCT 409
 985 GGGCGCTTCTGAGCTGAGAGATGAGCAGAGAGGCTCATGAGCAGAGCTGCTTGAACGA 1044

Db	410	-----	409
Qy	1045	GGATGAGTGTCAAGAAAGAGAGTCTCACGGGTAGAAAGCAGCCAGGGGTGGCGCA	1104
Db	410	-----	409
Qy	1105	CACGTGAATCCAGTACTTCGGGAGGCTGAGGCGAGAGAAATCGTTGAACCCGGAGGC	1164
Db	410	-----	409
Qy	1165	GGAGGTGTGGTGCAGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGA	1224
Db	410	-----GGTGGCAGCGACAGCTCATCAACCCGTTTGGAGAGATGATGACACTTTGA	460
Qy	1225	GACCAACTGGATTGTGCAGAGAAATTTGACAGGTGTCCCTTGTGGCTGTGGATGATGA	1284
Db	461	GACCAACTGGATTGTGCAGAGAGACTTGCAGAGTGTCCCTGTGGCTGTGATGATGCA	520
Qy	1285	CCAGAGCTGCTCGGATGAGACCGGACATGTACTGGAATTAAGCCGAGCCAGACGCC	1344
Db	521	CCACAGCTGCACCCATGGAGCAGAGATTTGATGGAATGATCCAGAACCAATCCCC	580
Qy	1345	CTACACAGTGTGTTCCGGCAGTTCCGTCAGAGCTCCTTTATGGGCTCCACTTCAACAT	1404
Db	581	CTACACAGTGTGTTTGTCCCAATGTCTCGCGACCCCTCCTTTTGGCTTCCACTTAAT	640
Qy	1405	CAGCCTGAACAAAGAGAGATGAGATTCCAGCCCAATCAGAGACAGAGAGATGTCA	1464
Db	641	CAGTCTGGGTAGGAGAGACATGGAATTCACAC-----GGAGAGAGAGAGAGAGGCGCA	694
Qy	1465	CGCTGGCAATCTGGCCGCTTCTTGGGCTCGAGTCCAGTCCATATCATCTCCACAGGC	1524
Db	695	CACGGCACTCTTGGCCACTTCTTGGGCTGCAATCCAGGACACACAGCCCCCAGAGC	754
Qy	1525	AAACTCAGAGACAAACTACTGTGGCCCAAGAGGAATCCTTCTCCACAGAGGCTGCC	1584
Db	755	AAACTCAAGAGAGAAACTCTGTGGCCCAAGAGAAAGAGCCATTTCCACGAGGGCCACCC	814
Qy	1585	CAAAACCCAGAGGACGCAACAGAGAGATTTGGGGCCAGAGAAAGCAACAAGGCTGGAA	1644
Db	815	CAAGAACCTCAGAGGGGGCCAGGCTGAGCTCTAGCCACAGAGAAAGACAGAGGCTGGAG	874
Qy	1645	GCTTAAAGGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAGAGGCCAGCTACTACAG	1704
Db	875	G-----GAGGGTGGCTTCAAGTCCGCTGCTCTGTGTGGAGGCGGGGTTCACACAG	925
Qy	1705	TGCCCCACAGAGCGCCCTCAGGCCCACTCCCAATGTTCTTCCCCCAATCAGACGCC	1764
Db	926	CGCCCCACAGACACCCCTCGGCCACACCCCTAATGTTCTTCCACCCGAGAGTACAGCCC	985
Qy	1765	GTCAAAGCTTCACAGTGTACAGGCGATAGA-----CACCAAGACAAAGCTTAAAGC	1818
Db	986	CTTAGGTCTTCGAGAGTCTCGGGGATATGATGAGCTGCGCAAGACACAGAGCTTCAACC	1045
Qy	1819	TGTGAGTTCGTGGGGCCCAAGAAAGTTTGAATTTGCTTCCTCAGAGAGAGATGGGGCTTGAT	1878
Db	1046	TGCGACATCCAGAGATCAAGAAAGTTTCAAGTTGCTCCAGAGAGGCGGAGGCTTCGCG	1105
Qy	1879	GGAGACACCCAGAGATATCTCAATGTGAGAGGAGAAACTGTGAGATTAACTCAGCGAATAT	1938
Db	1106	GGAGCCCCCTCAGAGGAGTCACTGTGAGAGGAAGACGTGTGAGATTAACTTGGCGAATCT	1165
Qy	1939	GCCAGAGATCCCGAAATATCACTCAAGAAACCTTTGGAACAATCAACCAACACATATCA	1998
Db	1166	GTCGAGAGGCCCGCA---CATCTCAAGAAAGACGAATTTGGAGCCACGATGGCATATCA	1222
Qy	1999	CACATACACTCAAGATCACAATGATATCTTATTTGGGCTTGAAGAACAGGAGTGAAGACA	2055
Db	1223	CGCTATCTCAAGACACACCGGATTCCTTACTGGGCTTGAAGAAACAGGATGAAGACACA	1282
Qy	2059	TTTCTTA 2064	

RESULT 5	AF139813/c	142092. bp	DNA	linear	PRI 22-APR-1999
LOCUS	AF139813				
DEFINITION	Homo sapiens clone pDJ759j12 chromosome 11 map 11q13, complete sequence.				
ACCESSION	AF139813				
VERSION	AF139813.1				
KEYWORDS	GI:4633837				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 142092) Petukhin, K., Kolst, M.J., Bakall, B., Li, W., Xie, G., Marknell, T., Sandgren, O., Foreman, K., Holmberg, G., Andersson, S., Vujic, M., Metzker, M.L., Caskey, C.T., and Wadelius, C.				
TITLE	Identification of the gene responsible for Best macular dystrophy				
JOURNAL	Nat. Genet. 19 (3), 241-247 (1998)				
MEDLINE	98324772				
PUBMED	9662395				
REFERENCE	2 (bases 1 to 142092) McGarry-Dugan, V.A., Hammond, H.A., Clement, M.K., Larson, D.R., Liu, X., Soderman, A.R., McGowan, J.M., DeAngelis, D.M., Lin, C., Fitzpatrick, E.S., Harrison, K.M., Petukhin, K., Caskey, C.T. and Metzker, M.L.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-APR-1999) Department of Human Genetics, Merck & Co., Inc., Sumneytown Pike, West Point, PA 19486, USA				
REMARK	IMPORTANT: This submission contains the entire insert of clone pDJ759j12 which comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. This sequence has been finished such that all consensus base calls consist of two or more separate clones with double-stranded coverage or two or more separate phred/Phrap calculated errors/10kb is 0.00.				
FEATURES	Location/Qualifiers				
source	1. 142092 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="11" /map="11q13" /clone="pDJ759j12" complement(19..1855) /rpt_family="L1" 1861..2001 /rpt_family="L1" 1990..2317 /rpt_family="Alu" complement(2031..2257) /rpt_family="SVA" 2320..2974 /rpt_family="L1" 2985..3540 /rpt_family="Alu" complement(4034..4124) /rpt_family="WIR" 4417..4995 /rpt_family="Alu" complement(4432..4936) /rpt_family="SVA" complement(5503..6111) /rpt_family="Alu" 5569..6095 /rpt_family="SVA" complement(7239..7542) /rpt_family="Alu"				

```
repeat_region 7306..7497
/rpt_family="SVA"
repeat_region 7752..7828
/rpt_family="MIR"
repeat_region 9099..9403
/rpt_family="Alu"
complement(9526..9897)
/rpt_family="MER1"
repeat_region 9635..9752
/rpt_family="MER1B"
repeat_region 11924..12197
/rpt_family="Alu"
complement(11955..12151)
/rpt_family="SVA"
repeat_region 16757..17039
/rpt_family="Alu"
complement(16775..17010)
/rpt_family="SVA"
repeat_region 18229..18485
/rpt_family="Alu"
complement(18220..18507)
/rpt_family="SVA"
repeat_region 20101..23446
/rpt_family="SVA"
/gene="HMFERRITH"
/note="Human ferritin heavy chain gene"
join(20101..20421,22218..22364,22621..22746,22842..23446)
/gene="HMFERRITH"
/product="ferritin heavy chain"
complement(21466..21731)
/rpt_family="Alu"
repeat_region 21506..21714
/rpt_family="SVA"
complement(23272..37375)
/gene="VMD2"
/note="Best macular dystrophy gene"
complement(join(23372..23629,24840..25478,27690..27841,
28155..28235,29427..29579,30274..30351,30740..30894,
31787..32020,32537..32631,35777..35964,37308..37375))
/gene="VMD2"
/product="Best macular dystrophy protein"
complement(23735..23954)
/rpt_family="Alu"
repeat_region 23789..23956
/rpt_family="SVA"
repeat_region 26429..26707
/rpt_family="Alu"
complement(26444..26696)
/rpt_family="SVA"
repeat_region 27103..27349
/rpt_family="Alu"
complement(27106..27323)
/rpt_family="SVA"
repeat_region 29161..29307
/rpt_family="Alu"
complement(29762..30043)
/rpt_family="Alu"
repeat_region 29786..30017
/rpt_family="SVA"
repeat_region 30462..30516
/rpt_family="MIR"
complement(31095..31409)
/rpt_family="Alu"
repeat_region 33005..33260
/rpt_family="Alu"
repeat_region 33247..33597
/rpt_family="LIMC2"
complement(33849..34431)
/rpt_family="Alu"
repeat_region 34076..34411
/rpt_family="SVA"
complement(34640..34890)
```

```
repeat_region /rpt_family="Alu"
34972..35401
/rpt_family="Alu"
repeat_region complement(34988..35392)
/rpt_family="SVA"
repeat_region 36238..36487
/rpt_family="Alu"
complement(36253..36457)
/rpt_family="SVA"
repeat_region complement(38661..38855)
/rpt_family="L1"
complement(40997..41224)
/rpt_family="Alu"
repeat_region 41056..41213
/rpt_family="SVA"
42745..43020
/rpt_family="Alu"
complement(42763..42966)
/rpt_family="SVA"
43312..43457
/rpt_family="Alu"
complement(43963..44217)
/rpt_family="L1"
44533)
/rpt_family="SVA"
complement(45300..45401)
/rpt_family="L1R5"
complement(45407..45533)
/rpt_family="L1R5"
complement(45542..45818)
/rpt_family="Alu"
45598..45770
/rpt_family="SVA"
complement(45862..46004)
/rpt_family="L1"
complement(45998..46282)
/rpt_family="Alu"
46022..46260
/rpt_family="SVA"
complement(46692..47583)
/rpt_family="Alu"
46885..47215
/rpt_family="SVA"
47363..47371
/rpt_family="SVA"
complement(47665..47773)
/rpt_family="Alu"
complement(47792..48077)
/rpt_family="Alu"
complement(48110..48388)
/rpt_family="L1"
complement(48827..49102)
/rpt_family="Alu"
48883..49110
/rpt_family="SVA"
complement(49395..49626)
/rpt_family="Alu"
50184..50437
/rpt_family="Alu"
```

Query Match 26.4%; Score 642.2; DB 9; Length 142092;
Best Local Similarity 99.5%; Pred. No. 1.3e-141; Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1405	CAGCTGAACAAAGAGATGAGTCCAGCCCAATCAGAGAGAGAGATGCTCA	1464
DB	25481	CAGCTGAACAAAGAGATGAGTCCAGCCCAATCAGAGAGAGAGATGCTCA	25422
QY	1465	CGCTGCAATCATTGCGCGCTTCTAGGCGCTGAGTCCATGATCACCATCTCCAGGCG	1524
DB	25421	CGCTGCAATCATTGCGCGCTTCTAGGCGCTGAGTCCATGATCACCATCTCCAGGCG	25362
QY	1525	AAATCAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCAGAGGCGCTGCC	1584

DB 25361 AAACCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAGAGGCGCTGCC 25302
 QY 1555 CAAAACCAAGGACCAAGCAAGAGCTTGGGGCCAGAGACACAAAGCCCTGGAA 1644
 DB 25301 CAAAACCAAGGACCAAGCAAGAGCTTGGGGCCAGAGACACAAAGCCCTGGAA 25242
 QY 1645 GCTTAAGGCTGTGAGGCGCTTCAAGTCTGAGGCGCCAGTATCAGAGGCGGCTACTACAG 1704
 DB 25241 GCTTAAGGCTGTGAGGCGCTTCAAGTCTGAGGCGCCAGTATCAGAGGCGGCTACTACAG 25182
 QY 1705 TGCCCCACAGAGCGCCCTCAGGCCCACTCCATGTTCTTCCCCCTAGAACATCAGCGCC 1764
 DB 25181 TGCCCCACAGAGCGCCCTCAGGCCCACTCCATGTTCTTCCCCCTAGAACATCAGCGCC 25122
 QY 1765 GTCAAGGCTTCAAGTCTCAGAGGCGATAGACACCAAGCAAAAGCTTAAAGACTGTGAG 1824
 DB 25121 GTCAAGGCTTCAAGTCTCAGAGGCGATAGACACCAAGCAAAAGCTTAAAGACTGTGAG 25062
 QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGGCGGAGTGGGGCCTTATGAGAGA 1884
 DB 25061 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGGCGGAGTGGGGCCTTATGAGAGA 25002
 QY 1885 CCCAGAACTATCTCAAGTGAAGAGGAAAGCTGTGAGTTTAACTGACGAGATATGCCAGA 1944
 DB 25001 CCCAGAACTATCTCAAGTGAAGAGGAAAGCTGTGAGTTTAACTGACGAGATATGCCAGA 24942
 QY 1945 GATCCCGGAAATATCACTCTCAAGAAACCTTTGGACATTCACCAACCAATACACACTAC 2004
 DB 24941 GATCCCGGAAATATCACTCTCAAGAAACCTTTGGACATTCACCAACCAATACACACTAC 24882
 QY 2005 ACTCAAGATCACAATGATGATCTTATTTGGGCGCTTGGAAACAGGAGT 2051
 DB 24881 ACTCAAGATCACAATGATGATCTTATTTGGGCGCTTGGAAACAGGAGT 24835

RESULT 6
 AC087451/c 163915 bp DNA linear HTG 04-JUL-2001
 LOCUS Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT
 DEFINITION
 AC087451
 AC087451.2 GI:14595836
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 163915)
 Birten, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 11, clone RP11-810P12
 Unpublished
 2 (bases 1 to 163915)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Birten, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barta, N., Bastien, V., Boguslavsky, L., Boukhaltier, B., Brown, A.,
 Camarata, J., Campoliano, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Glende, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Karatas, A., Laroque, K., Lamazares, R., Landers, T.,
 Lehoccky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P.,
 Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
 McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Plunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
 Roy, A., Santos, R., Schauer, S., Schuppback, R., Seaman, S., Severy, P.,
 Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Teeffey, S., Theodore, J.,
 Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,

TITLE
 JOURNAL
 COMMENT
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 4, 2001 this sequence version replaced g1:12039464.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Genome Center

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project information

Center project name: L11863
 Center clone name: 810_P_12

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 162458 bases at least Q40

Consensus quality: 162916 bases at least Q30

Consensus quality: 163173 bases at least Q20

Insert size: 170000; agarose-ff

Insert size: 163315; sum-of-ctrls

Quality coverage: 10.5 in Q20 bases; sum-of-ctrls

Quality coverage: 10.9 in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently

consists of 7 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 17285: contig of 17285 bp in length

17286 17385: gap of 100 bp

17386 19170: contig of 1795 bp in length

19171 19270: gap of 100 bp

19271 28600: contig of 9330 bp in length

28601 28700: gap of 100 bp

28701 47671: contig of 18971 bp in length

47672 47771: gap of 100 bp

47772 105063: contig of 57292 bp in length

105064 105163: gap of 100 bp

105164 142181: contig of 37018 bp in length

142182 142281: gap of 100 bp

142282 163915: contig of 21634 bp in length.

FEATURES

source

1. 163915

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-810P12"

/clone_lib="RPC1-11 Human Male BAC"

1. 17285

/note="assembly-fragment"

clone_end:SP6

vector_side:left"

17386. 19170

/note="assembly-fragment"

19271. 28600

/note="assembly-fragment"

28701. 47671

/note="assembly-fragment"

47772. 105063

/note="assembly-fragment"

105164. 142181

/note="assembly-fragment"

142282. 163915

/note="assembly-fragment"


```

clone_end:T7
vector_side:right
BASE COUNT 43950 a 38090 c 37945 g 43329 t 601 others
ORIGIN
Query Match 26.4%; Score 642.2; DB 2; Length 163915;
Best Local Similarity 99.5%; Pred. No. 1.4e-141;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1405 CAGCTGTAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 1464
    |||||||
DB 151430 CAGCTGTAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 151371
QY 1465 CGCTGCATCATTTGGCCGCTTCTAGGCGTCGAGTCCCATGATCCTCTCCAGAGGC 1524
    |||||||
DB 151370 CGCTGCATCATTTGGCCGCTTCTAGGCGTCGAGTCCCATGATCCTCTCCAGAGGC 151311
QY 1525 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCAGAGAGGCCTGCC 1584
    |||||||
DB 151310 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCAGAGAGGCCTGCC 151251
QY 1585 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
    |||||||
DB 151250 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151191
QY 1645 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAG 1704
    |||||||
DB 151190 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAG 151131
QY 1705 TGCCCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764
    |||||||
DB 151130 TGCCCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151071
QY 1765 GTCAGAGCTTCACAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1824
    |||||||
DB 151070 GTCAGAGCTTCACAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151011
QY 1825 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 1884
    |||||||
DB 151010 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 150951
QY 1885 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944
    |||||||
DB 150950 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150891
QY 1945 GATCCCCGGAATAATCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2004
    |||||||
DB 150890 GATCCCCGGAATAATCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150831
QY 2005 ACTCAAGATCAACATGATCTTATTGGGCTTGGAAACAGAGGATG 2051
    |||||||
DB 150830 ACTCAAGATCAACATGATCTTATTGGGCTTGGAAACAGAGGATG 150784

RESULT 7
AP003733 166867 bp DNA linear PRI 08-DEC-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-810P12,
DEFINITION complete sequence.
ACCESSION AP003733
VERSION AP003733.4 GI:17426128
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-810P12.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
AUTHORS Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 166867)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

```

```

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
JOURNAL Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.go.jp, URL:http://nyp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Dec 7, 2001 this sequence version replaced gi:17026124.

COMMENT
FEATURES
    source
        1. 166867
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q"
            /clone="RP11-810P12"
BASE COUNT 44705 a 38465 c 39268 g 44429 t
ORIGIN
Query Match 26.4%; Score 642.2; DB 9; Length 166867;
Best Local Similarity 99.5%; Pred. No. 1.4e-141;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1405 CAGCTGTAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 1464
    |||||||
DB 12486 CAGCTGTAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 12545
QY 1465 CGCTGCATCATTTGGCCGCTTCTAGGCGTCGAGTCCCATGATCCTCTCCAGAGGC 1524
    |||||||
DB 12545 CGCTGCATCATTTGGCCGCTTCTAGGCGTCGAGTCCCATGATCCTCTCCAGAGGC 12605
QY 1525 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCAGAGAGGCCTGCC 1584
    |||||||
DB 12606 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCAGAGAGGCCTGCC 12665
QY 1585 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
    |||||||
DB 12666 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12725
QY 1645 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAG 1704
    |||||||
DB 12726 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAG 12785
QY 1705 TGCCCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764
    |||||||
DB 12786 TGCCCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12845
QY 1765 GTCAGAGCTTCACAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1824
    |||||||
DB 12846 GTCAGAGCTTCACAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12905
QY 1825 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 1884
    |||||||
DB 12906 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAG 12965
QY 1885 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944
    |||||||
DB 12966 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13025
QY 1945 GATCCCCGGAATAATCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2004
    |||||||
DB 13026 GATCCCCGGAATAATCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13085
QY 2005 ACTCAAGATCAACATGATCTTATTGGGCTTGGAAACAGAGGATG 2051
    |||||||
DB 13086 ACTCAAGATCAACATGATCTTATTGGGCTTGGAAACAGAGGATG 13132

RESULT 8
AC004228 196080 bp DNA linear PRI 24-FEB-1999
LOCUS Homo sapiens Chromosome 11q12.2 pAC clone p0519013 containing
DEFINITION human gene, for ferritin heavy chain (FTH), complete sequence.
ACCESSION AC004228

```


VERSION AC004228.2 GI:4263838
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 196080)
AUTHORS Evans,G.A., Athanasidou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buetlner,J., Bumeister,R., Card,P., desaliboat,F., Dunn,J.,
English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
HTGS Submission
Unpublished

TITLE 2 (bases 1 to 196080)
JOURNAL Evans,G.A., Athanasidou,M., Basit,M., Bradbury,P., Brignac,S.,
Bumeister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
Hanson,S., Narayanaswamy,O., Newton,J., O'Brien,K., Patel,P.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Direct Submission

TITLE Submitted (26-FEB-1998) Genome Science & Technology Center,
JOURNAL University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 196080)
REFERENCE Evans,G.A., Athanasidou,M., Aguayo,P., Armstrong,D., Basit,M.,
AUTHORS Buetlner,J., Butler,C., Card,P., desaliboat,F., Dunn,J.,
English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T.
Direct Submission

TITLE Submitted (24-FEB-1999) Genome Science & Technology Center,
JOURNAL University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
On Feb 24, 1999 this sequence version replaced gi:2911733.
COMMENT IMPORTANT: This submission contains the entire insert of clone
pJ519013. pJ519013 comes from the RPI-3 PAC library constructed
at the Roswell Park Cancer Institute by the Pletier de Jong group.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
11p12.2 Best's disease region mapped between STS D11S461 and EST
A1NAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA), STS
D11S699 and WI-7524
MAPPED CLONE OVERLAP: PACs pJ466a11 and pJ0756b9.

FEATURES
source location/Qualifiers
1..196080
/organism="Homo sapiens"
/db_xref="taxon:9606"
complement(2210..2482)
repeat_region
/rpt_family="Alu"
3457..3743
/rpt_family="Alu"
complement(20145..20758)
repeat_region
/rpt_family="Alu"
complement(23194..23599)
repeat_region
/rpt_family="Alu"
complement(227367..27685)
repeat_region
/rpt_family="Alu"
28047..28457
/rpt_family="Alu"
complement(29390..29565)
repeat_region
/rpt_family="Alu"
complement(31076..31255)
repeat_region
/rpt_family="Alu"
complement(34948..35045)
repeat_region
/rpt_family="Alu"
35232..35538
complement(36242..36540)
repeat_region
/rpt_family="Alu"

repeat_region 37238..37522
/rpt_family="Alu"
complement(38555..38650)
repeat_region
/rpt_family="Alu"
complement(40483..41164)
repeat_region
/rpt_family="Alu"
42042..42116
/rpt_family="MIR"
repeat_region
/rpt_family="LRR12"
42677..42817
/rpt_family="LRR12"
42856..43157
/rpt_family="Alu"
43148..43360
/rpt_family="LRR12"
complement(43734..43999)
repeat_region
/rpt_family="LRR12"
complement(46998..47169)
repeat_region
/rpt_family="MIR"
complement(48048..48346)
repeat_region
/rpt_family="Alu"
48482..48737
/rpt_family="Alu"
complement(48893..49037)
repeat_region
/rpt_family="L1"
49118..49292
/rpt_family="Alu"
49791..50059
/rpt_family="Alu"
50436..50501
/rpt_family="Alu"
complement(51147..51441)
repeat_region
/rpt_family="Alu"
complement(52700..52979)
repeat_region
/rpt_family="Alu"
53281..53568
/rpt_family="Alu"
54071..54204
/rpt_family="Alu"
54506..54793
/rpt_family="Alu"
54807..55061
/rpt_family="Alu"
complement(64533..64826)
repeat_region
/rpt_family="Alu"
complement(65245..65783)
repeat_region
/rpt_family="Alu"
complement(68465..68732)
repeat_region
/rpt_family="Alu"
complement(70191..70476)
repeat_region
/rpt_family="Alu"
70560..70692
/rpt_family="Alu"
72294..72470
/rpt_family="Alu"
complement(72800..73095)
repeat_region
/rpt_family="Alu"
73424..73695
/rpt_family="Alu"
73784..74071
/rpt_family="Alu"
74652..74935
/rpt_family="Alu"
complement(75265..75514)
repeat_region
/rpt_family="Alu"
complement(76086..76353)
repeat_region
/rpt_family="Alu"
complement(78038..78312)
repeat_region
/rpt_family="Alu"
78513..78734
/rpt_family="Alu"
78866..79180
/rpt_family="Alu"
complement(79384..79673)
repeat_region

```
repeat_region /rpt_family="Alu" .81690)
/rpt_family="Alu"
repeat_region /rpt_family="Alu" .82288
/rpt_family="Alu" .83600)
/rpt_family="MIR" .87624)
/rpt_family="Alu" .87987)
/rpt_family="Alu" .88186)
/rpt_family="Alu" .88472)
/rpt_family="Alu" .88822)
/rpt_family="Alu" .89049)
/rpt_family="Alu" .89308)
/rpt_family="MER20" .89729)
/rpt_family="Alu" .90487)
/rpt_family="Alu" .90504)
/rpt_family="Alu" .90618)
/rpt_family="Alu" .93313)
/rpt_family="Alu" .93952)
/rpt_family="Alu" .96938)
/rpt_family="Alu" .97517)
/rpt_family="Alu" .104514)
/rpt_family="Alu" .104519)
/rpt_family="Alu" .109738)
/rpt_family="Alu" .109979)
/rpt_family="Alu" .110022)
/rpt_family="Alu" .110141)
/rpt_family="Alu" .110216)
/rpt_family="Alu" .111091)
/rpt_family="Alu" .113210)
/rpt_family="Alu" .113419)
/rpt_family="I1" .115005)
/rpt_family="MIR" .115685)
/rpt_family="MIR" .115791)
/rpt_family="MIR" .116870)
/rpt_family="MIR" .117005)
/rpt_family="MIR" .117877)
/rpt_family="Alu" .118172)
/rpt_family="Alu" .118352)
/rpt_family="Alu" .118600)
/rpt_family="Alu" .118679)
/rpt_family="Alu" .118813)
/rpt_family="Alu" .121233)
/rpt_family="Alu" .121520)
/rpt_family="Alu" .122668)
/rpt_family="Alu" .123071)
```

Query Match	26.4%	Score 642.1	DB 9	Length 196080
Best Local Similarity	99.5%	Pred. No. 1.4e-141		
Matches 644	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY	1405	CAGCGTGAACAAGAGAGATGAGACTTCACGCCCAATCAGAGAGAGAGAGATGCTCA	1464	
Db	191325	CAGCGTGAACAAGAGAGAGATGAGACTTCAGGCCCAATCAGAGAGAGAGAGATGCTCA	191384	
QY	1465	CGCTGGCATCATTTGGCCGCTTCTTAGGCTGCAAGTCCCATGATCACCATCTCTCCAGGGC	1524	
Db	191385	CGCTGGCATCATTTGGCCGCTTCTTAGGCTGCAAGTCCCATGATCACCATCTCTCCAGGGC	191444	
QY	1525	AAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGGAATCCCTTCACAGAGGGCCGCC	1584	
Db	191445	AAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGGAATCCCTTCACAGAGGGCCGCC	191504	
QY	1585	CAAAACCCACAGGCGCCCAACAGAACGTTAGGGGCCAGGAAGACAAACAGGCCCTGGAA	1644	
Db	191505	CAAAACCCACAGGCGCCCAACAGAACGTTAGGGGCCAGGAAGACAAACAGGCCCTGGAA	191564	
QY	1645	GCTTAAGGCTGTGACGCGCTTCAAGTCTTGCCCACTGTATCAGAGGCCAGGCTACTACAG	1704	

Db	191565	GCTTAAAGGCTGTGGACGGCTTCAGTGCTGC	CCCACTGTATCAGAGGCGCAGGCTACTACAG	1916234
OY	1705	TGCCCCACAGACGGCCCCCTCACC	CCCACACTGCCATGTCTTCCCCCTAGAACCATACAGGCC	1764
Db	191625	TGCCCCACAGACGGCCCCCTCACC	CCCACACTGTCTTCCCCCTAGAACCATACAGGCC	1916844
OY	1765	GTCAAACCTTCACAGTGTACAGGCGATAGAC	ACCACAAGCAAAGAGTTTAAGACTGTGAG	1824
Db	191685	GTCAAAGCTTCACAGTGTACAGGCGATAGAC	ACCACAAGCAAAGAGTTTAAGACTGTGAG	1917444
OY	1825	TTCTGGGGGCCAAGAAAAGTTTTGAATTGCTCT	CACAGACCGCATGGGGCCTTGATGAGCA	1884
Db	191745	TTCTGGGGGCCAAGAAAAGTTTTGAATTGCTCT	CACAGACCGCATGGGGCCTTGATGAGCA	1918044
OY	1885	CCCGAAGATATCTCAAGTGAGGAGAGAAAAC	TGGAGTTAACCTGACGATATGCGACAG	1944
Db	191805	CCCGAAGATATCTCAAGTGAGGAGAGAAAAC	TGGAGTTAACCTGACGATATGCGACAG	1918664
OY	1945	GATCCCGGAAATCACCTCAAGAACCTTTGGAC	AATCACCAACCATACACACTAC	2004
Db	191865	GATCCCGGAAATCACCTCAAGAACCTTTGGAC	AATCACCAACCATACACACTAC	1919244
OY	2005	ACTCAAGATACATGATGATCCTTATTGGGCTT	TGGAAAAACAGGATG	2051
Db	191925	ACTCAAGATACATGATGATCCTTATTGGGCTT	TGGAAAAACAGGATG	191971
RESULT 9				
HSVMD2P10				
LOCUS	HSVMD2P10	706 bp	DNA	linear PRI 22-OCT-1998
DEFINITION	Homo sapiens vitelliform macular dystrophy protein (VMD2) gene,			
	exon 10.			
ACCESSION	AF073499			
VERSION	AF073499.1	GI:3598873		
KEYWORDS				
SEGMENT				
SOURCE				
ORGANISM	Homo sapiens.			
	Homo sapiens.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 706)			
AUTHORS	Stohr,H., Marguardt,A., Rivera,A., Cooper,P.R., Nowak,N.J., Shows,T.B., Gerhard,D.S. and Weber,B.H.			
TITLE	A gene map of the Best's vitelliform macular dystrophy region in chromosome 11q12-q13.1			
JOURNAL	Chromosome Res. 8 (1), 48-56 (1998)			
MEDLINE	98112782			
PUBMED	9445487			
REFERENCE	2 (bases 1 to 706)			
AUTHORS	Marguardt,A., Stohr,H., Passmore,L.A., Kramer,F., Rivera,A. and Weber,B.H.			
TITLE	Mutations in a novel gene, VMD2, encoding a protein of unknown properties cause juvenile-onset vitelliform macular dystrophy (Best's disease)			
JOURNAL	Hum. Mol. Genet. 7 (9), 1517-1525 (1998)			
MEDLINE	98367043			
PUBMED	9700209			
REFERENCE	3 (bases 1 to 706)			
AUTHORS	Marquardt,A., Stohr,H., Passmore,L., Kraemer,F., Rivera,A. and Weber,B.H.F.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-JUN-1998) Human Genetics, University, Bionzentrum, Am Hubland, Wurzburg 97074, Germany			
FEATURES				
SOURCE				
	1..706			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="11"			
	/map="11q12-q13.1"			
	/clone="PAC d1466a11"			
	36..674			
	/gene="VMD2"			
	/number=10			
exon				

BASE COUNT 204 a 209 c 164 g 129 t

ORIGIN

Query Match 26.4% Score 640.6; DB 9; Length 706;
 Best Local Similarity 99.4%; Pred. No. 1.3e-141;
 Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1405 CAGCCTTAACAAAGAGAGATGAGTTCACGCCCAATCAGACAGACGAGAGATGCTCA 1464
 DB 33 CAGCCTTAACAAAGAGAGATGAGTTCACGCCCAATCAGACAGACGAGAGATGCTCA 92

QY 1465 CGCGGATCATTTGGCGCTTCTTAGGGCTGAGTCCATATCCATCTCCAGAGGC 1524
 DB 93 CGCGGATCATTTGGCGCTTCTTAGGGCTGAGTCCATATCCATCTCCAGAGGC 152

QY 1525 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTTCACGAGAGGCTCC 1584
 DB 153 AAATCAGAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTTCACGAGAGGCTCC 212

QY 1585 CAAAACCCACAGGACGACCAACAGACGTTAGGGCCGAGAGACAAACAGGCTGGAA 1644
 DB 213 CAAAACCCACAGGACGACCAACAGACGTTAGGGCCGAGAGACAAACAGGCTGGAA 272

QY 1645 GCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGATCAGAGGCGGCTACTAG 1704
 DB 273 GCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGATCAGAGGCGGCTACTAG 332

QY 1705 TGGCCACAGACAGCGCCCTCAGCCCACTCCATGCTTCTTCCCTAGAACCATCAGGCC 1764
 DB 333 TGGCCACAGACAGCGCCCTCAGCCCACTCCATGCTTCTTCCCTAGAACCATCAGGCC 392

QY 1765 GTCAAACCTTCACAGTGTACAGGATGACACCAACAAAGCTTAAAGACTGTAG 1824
 DB 393 GTCAAACCTTCACAGTGTACAGGATGACACCAACAAAGCTTAAAGACTGTAG 452

QY 1825 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGATGAGCA 1884
 DB 453 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGATGAGCA 512

QY 1885 CCCAGAGATATCTCAAGTGAAGAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 1944
 DB 513 CCCAGAGATATCTCAAGTGAAGAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 572

QY 1945 GATCCCCGGAATAATCACTCAAGAACTTTGGACATACCAACCAATACACATAC 2004
 DB 573 GATCCCCGGAATAATCACTCAAGAACTTTGGACATACCAACCAATACACATAC 632

QY 2005 ACTCAAGATCACATGATCTTATTTGGGCTTGGAACAGGGATG 2051
 DB 633 ACTCAAGATCACATGATCTTATTTGGGCTTGGAACAGGGATG 679

RESULT 10
 AC003025 112309 bp DNA linear PRI 23-JUL-1998
 LOCUS Human Chromosome 11p12.2 PAC clone PDJ466a11, complete sequence.
 AC003025
 AC003025.1 GI:33737308
 DEFINITION
 AC003025.1
 VERSION
 AC003025.1
 KEYWORDS
 HTG.
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eultheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 Evans,G.A., Athanaslou,M., Bradbury,P., Brignac,S., Bumeister,R.,
 1 (bases 1 to 112309)
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Buetner,J., Bumeister,R., Card,P., desaliboat,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
 HTGS Submission
 Unpublished
 JOURNAL
 2 (bases 1 to 112309)

AUTHORS
 Evans,G.A., Athanaslou,M., Bradbury,P., Brignac,S., Bumeister,R.,
 Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M.,
 Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S.,
 Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P.,
 Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R.,
 Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
 Direct Submission

TITLE
 Submitted (21-OCT-1997) Genome Science and Technology Center,
 University of Texas Southwestern Medical Center at Dallas, 5323
 Harry Hines Blvd., Dallas, TX 75235-8591, USA
 3 (bases 1 to 112309)

REFERENCE
 AUTHORS
 Evans,G.A., Athanaslou,M., Bradbury,P., Brignac,S., Bumeister,R.,
 Buetner,J., Bumeister,R., Card,P., desaliboat,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
 Direct Submission

TITLE
 Submitted (23-JUL-1998) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines
 Blvd., Dallas, TX 75235-8591, USA

COMMENT
 IMPORTANT: This submission contains the entire insert of clone
 PDJ466a11. PDJ466a11 comes from a PAC library constructed at the
 Roswell Park Cancer Institute by the Pieter de Jong group. This
 clone has been finished according to strict quality criteria and
 attempts have been made to resolve all base calling problems such
 as compressions and repetitive elements. The expected Phred/Phrap
 calculated errors/10kb is 0.18. In addition, this sequence has
 been finished such that 99.9% of consensus base calls consist of
 either double-stranded coverage or 2 types of labeling chemistry on
 one strand.
 Further information regarding the map of this region or
 annotation of PDJ466a11 can be found at
 http://gsstec.smed.edu/chromosome1.htm.
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
 11p12.2 Best's disease region mapped between SRS D11S461 and EST
 AHNK. This region spans over 1.5 Mbp.
 MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA)
 MAPPED CLONE OVERLAP: HTGS submitted PAC clones PDJ519013 and
 PDJ756b9.

FEATURES
 source
 1. 112309
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 complement(4341..4656)
 repeat_region
 /rpt_family="Alu"
 complement(9880..10120)
 repeat_region
 /rpt_family="Alu"
 10163..10282
 /rpt_family="Alu"
 complement(11232..
 /rpt_family="Alu"
 13351..13560
 /rpt_family="Alu"
 complement(15092..
 /rpt_family="MIR"
 15826..15932
 /rpt_family="MIR"
 17011..17146
 /rpt_family="MIR"
 /rpt_family="MIR"
 complement(18018..
 /rpt_family="Alu"
 18493..18741
 /rpt_family="Alu"
 18820..18954
 /rpt_family="Alu"
 complement(21374..
 /rpt_family="Alu"
 22809..23212
 /rpt_family="Alu"
 23499..23752
 /rpt_family="Alu"

repeat_region 24118..24472 /rpt_family="Alu"
 repeat_region 25005..25113 /rpt_family="MIR"
 repeat_region 31160..31442 /rpt_family="Alu"
 repeat_region complement(32651..32947) /rpt_family="Alu"
 repeat_region complement(35970..36025) /rpt_family="MIR"
 repeat_region 38958..39145 /rpt_family="Alu"
 repeat_region 44878..45480 /rpt_family="Alu"
 repeat_region 49769..50071 /rpt_family="Alu"
 repeat_region complement(50172..50310) /rpt_family="Alu"
 repeat_region 50902..50982 /rpt_family="MIR"
 repeat_region 51038..51659 /rpt_family="Alu"
 repeat_region 54494..54769 /rpt_family="Alu"
 repeat_region 55529..55623 /rpt_family="THER1"
 repeat_region complement(55691..55967) /rpt_family="Alu"
 repeat_region complement(56049..56257) /rpt_family="Alu"
 repeat_region 56721..57100 /rpt_family="THER1"
 repeat_region 58402..58680 /rpt_family="Alu"
 repeat_region complement(58748..59345) /rpt_family="Alu"
 repeat_region 59678..59904 /rpt_family="Alu"
 repeat_region 61149..61424 /rpt_family="Alu"
 repeat_region complement(61930..62242) /rpt_family="Alu"
 repeat_region 62276..62710 /rpt_family="MIR1"
 repeat_region 62915..63118 /rpt_family="Alu"
 repeat_region complement(63349..63455) /rpt_family="Alu"
 repeat_region 64339..64614 /rpt_family="Alu"
 repeat_region complement(64707..64984) /rpt_family="Alu"
 repeat_region complement(65596..65811) /rpt_family="Alu"
 repeat_region complement(66500..66753) /rpt_family="Alu"
 repeat_region 67311..67542 /rpt_family="Alu"
 repeat_region 67835..68119 /rpt_family="Alu"
 repeat_region 68860..69150 /rpt_family="Alu"
 repeat_region 69164..69272 /rpt_family="Alu"
 repeat_region 69354..70245 /rpt_family="Alu"
 repeat_region 70655..70939 /rpt_family="Alu"
 repeat_region 70933..71075 /rpt_family="L1"
 repeat_region 71119..71395 /rpt_family="Alu"
 repeat_region 71404..71530 /rpt_family="Alu"

repeat_region /rpt_family="LFR5"
 repeat_region 71536..71637 /rpt_family="LFR5"
 repeat_region 72720..72948 /rpt_family="L1"
 repeat_region complement(73488..73633) /rpt_family="Alu"
 repeat_region complement(73925..74200) /rpt_family="Alu"
 repeat_region 75721..75955 /rpt_family="Alu"
 repeat_region 77185..77449 /rpt_family="Alu"
 repeat_region 78093..78222 /rpt_family="L1"
 repeat_region 78227..78294 /rpt_family="L1"
 repeat_region complement(80458..80707) /rpt_family="Alu"
 repeat_region complement(81544..81974) /rpt_family="Alu"
 repeat_region 82055..82305 /rpt_family="Alu"
 repeat_region 82514..83096 /rpt_family="Alu"
 repeat_region complement(83685..83940) /rpt_family="Alu"
 repeat_region 85336..85850 /rpt_family="Alu"
 repeat_region complement(86429..86483) /rpt_family="MIR"
 repeat_region 86902..87183 /rpt_family="Alu"
 repeat_region 87646..87792 /rpt_family="Alu"

Query Match 26.4%; Score 640.6; DB 9; Length 112309;
 Best Local Similarity 99.4%; Pred. No. 3.1e-141;
 Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1405 CAGCGTGAACAAAGAGAGATGAGTTCACAGCCCAATCAGAGAGAGAGATGCTCA 1464
 DB 91470 CAGCGTGAACAAAGAGAGATGAGTTCACAGCCCAATCAGAGAGAGATGCTCA 91529
 QY 1465 CGCTGGCATCTTGGCCGCTTCCAGGCGTGCAGTCCCATGATCCTCCAGGCG 1524
 DB 91530 CGCTGGCATCTTGGCCGCTTCCAGGCGTGCAGTCCCATGATCCTCCAGGCG 91589
 QY 1525 AAACCTCAAGCACAACCTACTGTGGCCCAAGAGGAAATCCCTTCCACGAGGCGTCC 1584
 DB 91590 AAACCTCAAGCACAACCTACTGTGGCCCAAGAGGAAATCCCTTCCACGAGGCGTCC 91649
 QY 1585 CAAAACCCACAGGCGCCAAACAGAAAGTTAGGGCCAGAGACAAAGCGCTGGAA 1644
 DB 91650 CAAAACCCACAGGCGCCAAACAGAAAGTTAGGGCCAGAGACAAAGCGCTGGAA 91709
 QY 1645 GCTTAAGCGTGTGAGCGCTTCAAGTCTGGGCCACTGATTCAGAGGCGAGCTACTACAG 1704
 DB 91710 GCTTAAGCGTGTGAGCGCTTCAAGTCTGGGCCACTGATTCAGAGGCGAGCTACTACAG 91769
 QY 1705 TGCCCAAGAGAGCGCCCTCAGCCCTCAGCTCCATGTTCTCCCTAGAACCTATGAGGCC 1764
 DB 91770 TGCCCAAGAGAGCGCCCTCAGCCCTCAGCTCCATGTTCTCCCTAGAACCTATGAGGCC 91829
 QY 1765 GTCAAAAGCTTCACAGTGTACAGGCGATAGACCAAAAGCTTAAAGACTGTGAG 1824
 DB 91830 GTCAAAAGCTTCACAGTGTACAGGCGATAGACCAAAAGCTTAAAGACTGTGAG 91889
 QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTGAGAGGAGGAGGCGCTGATGAGCA 1884
 DB 91890 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTGAGAGGAGGAGGCGCTGATGAGCA 91949
 QY 1885 CCCAAGATATCTCAAGTGAAGAGAGAAACTGTGAGATTAACTGACGATATGCGAGA 1944

|||||

Db 91950 CCAGAGATATCTCAAGTGAAGAGAAATCTGTGAGTTTAACTGACGGATATGCCGGA 92009

Qy 1945 GATCCCGGAAATACCTCAAGAACCTTTGGAACAATCACCACCAACACATACACTAC 2004

Db 92010 GATCCCGGAAATACCTCAAGAACCTTTGGAACAATCACCACCAACACATACACTAC 92069

Qy 2005 ACTCAAGATCAGATGATGCTTATTTGGGCTTGGAACAAGGATG 2051

Db 92070 ACTCAAGATCAGATGATGCTTATTTGGGCTTGGAACAAGGATG 92116

RESULT 11

AC084857 133683 bp DNA linear HTG 22-NOV-2000

LOCUS

DEFINITION Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT

SEQUENCE 13 unordered pieces.

AC084857

AC084857.1 GI:11276215

HTG: HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE

1 (bases 1 to 133683)

Blumen, B., Linton, L., Nusbaum, C., Lander, E., Abrahams, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L., Boukhalil, B., Brown, A., Burkett, G., Campopiano, A., Castelle, A., Choe, P., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heath, F., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, C., Liu, G., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meltzer, J., Meneses, L., Miya, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisanti, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Soung, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tefaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A., and Zody, M.

Direct Submission

Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

----- Project Information

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L11038

Center clone name: 3231_N5

----- Summary Statistics

Sequencing vector: Plasmid: n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 126761 bases at least Q40

Consensus quality: 129952 bases at least Q30

Consensus quality: 131351 bases at least Q20

Insert size: 12800; agarose-fp

Insert size: 132483; sum-of-contigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1459: contig of 1459 bp in length

1460 1559: gap of 100 bp

1560 2520: contig of 961 bp in length

2521 2620: gap of 100 bp

2621 6334: contig of 3714 bp in length

6335 6434: gap of 100 bp

6435 11913: contig of 5479 bp in length

11914 12013: gap of 100 bp

12014 20756: contig of 8743 bp in length

20757 20856: gap of 100 bp

20857 29735: contig of 8879 bp in length

29736 29835: gap of 100 bp

29836 38645: contig of 8810 bp in length

38646 38745: gap of 100 bp

38746 52193: contig of 13448 bp in length

52194 52293: gap of 100 bp

52294 65209: contig of 12916 bp in length

65310 79796: contig of 14487 bp in length

79797 79896: gap of 100 bp

79897 95896: contig of 16000 bp in length

95897 95996: gap of 100 bp

95997 113419: contig of 17423 bp in length

113420 113519: gap of 100 bp

113520 133683: contig of 20164 bp in length.

Location/Qualifiers

1. 133683

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-3231N5"

/clone_11b="CITD Human BAC"

1. 1459

/note="assembly-fragment"

1560. 2520

/note="assembly-fragment"

2621. 6334

/note="assembly-fragment"

6435. 11913

/note="assembly-fragment"

12014. 20756

/note="assembly-fragment"

20857. 29735

/note="assembly-fragment"

29836. 38645

/note="assembly-fragment"

38746. 52193

/note="assembly-fragment"

52294. 65209

/note="assembly-fragment"

65310. 79796

/note="assembly-fragment"

79897. 95896

/note="assembly-fragment"

95997. 113419

/note="assembly-fragment"

113520. 133683

/note="assembly-fragment"

BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others

ORIGIN

Query Match 26.4%; Score 640.6; DB 2; Length 133683;

FEATURES * 142130 160169: contig of 18040 bp in length.
Location/Qualifiers
source
1. 160169
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-810P12"
1. 1926
misc_feature
/note="assembly_name:Contig18"
2027. 4202
/note="assembly_name:Contig19"
4303. 7163
/note="assembly_name:Contig20"
vector_end:77
vector_side:left"
7264. 9961
misc_feature
/note="assembly_name:Contig21"
1062. 12575
/note="assembly_name:Contig22"
12676. 15545
misc_feature
/note="assembly_name:Contig23"
15646. 20941
/note="assembly_name:Contig24"
21042. 23966
misc_feature
/note="assembly_name:Contig25"
24067. 27231
/note="assembly_name:Contig26"
27332. 30550
misc_feature
/note="assembly_name:Contig27"
30651. 35202
/note="assembly_name:Contig28"
35303. 40201
misc_feature
/note="assembly_name:Contig29"
40302. 45758
/note="assembly_name:Contig30"
45859. 51249
misc_feature
/note="assembly_name:Contig31"
51350. 58267
misc_feature
/note="assembly_name:Contig32"
58368. 66916
misc_feature
/note="assembly_name:Contig33"
67017. 77145
misc_feature
/note="assembly_name:Contig34"
77246. 86493
misc_feature
/note="assembly_name:Contig35"
86594. 96688
misc_feature
/note="assembly_name:Contig36"
96789. 107169
misc_feature
/note="assembly_name:Contig37"
107270. 117047
misc_feature
/note="assembly_name:Contig38"
117148. 127120
misc_feature
/note="assembly_name:Contig39"
vector_end:sp6
vector_side:right"
12721. 142029
misc_feature
/note="assembly_name:Contig40"
142130. 160169
/note="assembly_name:Contig41"
BASE COUNT 41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN
Query Match 22.88; Score 552.8; DB 2; Length 160169;
Best Local Similarity 99.3%; Pred. No. 2.5e-120;
Matches 576; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1405 CAGCGTAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 1464
DB 67594 CAGCGTAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATG-TCA 67536
QY 1465 CGCTGGCATCTTGGCGCTTCCTGAGCTGACATCCAGATCACCATCTCCAGGGC 1524
DB 67535 CGCTGGCATCTTGGCGCTTCCTGAGCTGACATCCAGATCACCATCTCCAGGGC 67476

QY 1525 AAATCAAGAGCCAAATCTGTGGCCCAAGAGGGAATCCCTTCACAGAGGCGCTGCC 1584
DB 67475 AAATCAAGAGCCAAATCTGTGGCCCAAGAGGGAATCCCTTCACAGA-66CCTGCC 67417
QY 1585 CAAAACCCCAAGAGGCGCCAAACAGAAAGTTAGGCGCCAGAGAGACAAAGGCTGGA 1644
DB 67416 CAAAACCCCAAGAGGCGCCAAACAGAAAGTTAGGCGCCAGAGAGACAAAGGCTGGA 67357
QY 1645 GCTTAAAGGCTGTGAGAGCCCTTCAAGTCTGGCCCACTGTATCAGAGCCAGGCTACTACAG 1704
DB 67356 GCTTAAAGGCTGTGAGAGCCCTTCAAGTCTGGCCCACTGTATCAGAGCCAGGCTACTACAG 67297
QY 1705 TGCCCCACAGAGCGCCCTCAGCCCACTCCCAAGTTCTCCCTGAACATCAGAGCC 1764
DB 67296 TGCCCCACAGAGCGCCCTCAGCCCACTCCCAAGTTCTCCCTGAACATCAGAGCC 67237
QY 1765 GTCAAAGCTTCACAGTGTCAAGCATATGACACCAAGAGCAAAAGCTTAAAGCTGTAG 1824
DB 67236 GTCAAAGCTTCACAGTGTCAAGCATATGACACCAAGAGCAAAAGCTTAAAGCTGTAG 67177
QY 1825 TTCTGGGGCCCAAGAAAGTTTGATTTGCTCTCTCAGAGAGGATGGGCTTGATGAGCA 1884
DB 67176 TTCTGGGGCCCAAGAAAGTTTGATTTGCTCTCTCAGAGAGGATGGGCTTGATGAGCA 67117
QY 1885 CCCAAGATATCTCAAGTGTGAGGAGAAACGTGTGAGTTTAACTGACGATATGCCAGA 1944
DB 67116 CCCAAGATATCTCAAGTGTGAGGAGAAACGTGTGAGTTTAACTGACGATATGCCAGA 67057
QY 1945 GATCCCCGAAATCACTCAAGAAACCTTTGGAAACAATCA 1944
DB 67056 GATCCCCGAAATCACTCAAGAAACCTTTGGAAACAATCA 67017

RESULT 13
AF440756 1908 bp mRNA linear 12-JUL-2002
LOCUS AF440756
DEFINITION Homo sapiens vitelliform macular dystrophy 2-like protein 1 mRNA,
complete cds.
ACCESSION AF440756
VERSION AF440756.1 GI:21734839
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Marguardt, A., Stohr, H., Passmore, L.A., Kramer, F., Rivera, A. and
Weber, B.H.
TITLES Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(Best's disease)
Hum. Mol. Genet. 7 (9), 1517-1525 (1998)
JOURNAL MEDLINE
PUBMED 98367043
REFERENCE
AUTHORS Stohr, H., Marguardt, A., Nanda, I., Schmid, M. and Weber, B.H.
TITLES Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family
Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
JOURNAL MEDLINE
PUBMED 12032738
REFERENCE
AUTHORS Stohr, H., Marguardt, A. and Weber, B.H.F.
TITLES Three novel human VMD2-like protein genes are members of the
evolutionary highly conserved RFP family
Unpublished
4 (bases 1 to 1908)
Stohr, H., Marguardt, A. and Weber, B.H.F.
Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg,
Biozentrum, Am Hubland, Wuerzburg 97074, Germany
Location/Qualifiers

source
1. 1908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.12-p13.2"
1.1530
/note="VMD2L1: putative membrane protein; Pfam01062;
worm_family:8
/codon_start=1
/product="vitellogenin macular dystrophy 2-like protein 1"
/protein_id="AA076995.1"
/db_xref="GI:21734840"
/translation="MTVTYARVANARFSGESOLLILNRGSIYLKMLRELLCETGFTY
ALSAAYRPLTSGKRFYKLYICDOVASLIPVSLFVYTLVVRNWSQYLCML
PALMCVAVGTHGRDNRRLYRPLIRAGLSAVLLRSVTAVERPDIHAYEA
GPMREERKREKFNLSINKYPCWWSNLAOARRGRIIDNSALKLELVEFR
GKGMFLHYDMISVPLVYQVVTALYSTFLACLIDRNFQVSLVDEKDYDLAYEK
TLIOFFEYAGMLKVAEQLINPGEDEDDDEFNLIDRNFQVSLVDEKDYDLAYEK
DLVMDAERAPYATATVQLRPSFQSGSTFDITLAKEDMQRDLGDEPMEAPGD
FLDRLIPAGAGVAGPLGRRLSFLRKNSCVSEASTGSCSAVVEGADECSGD
PLDPLPEPBPAPGPPPLILPEPVPFSLVMPGPRGAPPLPSPIGEEBENT
A"

BASE COUNT 337 a 604 c 551 g 416 t

ORIGIN

Query Match 17.7%; Score 431; DB 9; Length 1908;
Best Local Similarity 68.4%; Pred. No. 1.1e-21;
Matches 596; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

105 ATGACATCACTACACAGCCAGAGTGGCTTAATGCCGCTTACGCTCCCGCTG 164
1 ATGACGCTACACACAGCCAGAGTGGCTTAATGCCGCTTACGCTCCCGCTG 60
165 CTGCTGCTGCGCGGCGGAGCATCTACAGCTGCTATATGAGCGAGTCTTAATCTTCG 224
61 CTGCTACTGCTGCGGAGCATCTACAACTCTGCGGAGAGTCTGCTGCTCTT 120
225 CTGCTACTGCTGCGGAGCATCTGCTTATTAATGCTGCGGCGGAGAGAGCAACACTG 284
121 GGGTCTCATGCGGCTGAGTGTGCTGACCGCTTGTGCTGACCGAGGCGAGAGCGC 180
285 ATGTTGGAACACTGCTGATGCGAGAGCTATCCAGCTATCCCATTTCTTC 344
181 TACTTCGAGAGCTGATGATTTATGTAACAGTATGCCAGCTATCCGCTGCTTC 240
345 GGTGCGGCTCTACGCTGCTGCTGAGCCGCTGAGCCGCTGAGCAACAGTACGAGACTG 404
241 GGTGCGGCTTATGATGAGCTGCTGAGCAACCGCTGAGCCGCTGAGCAACAGTATGACTG 300
405 CCGTGGCCGAGCGGCTCATGAGCTGCTGCGGCTTCGTCGAGAGGAGAGGAGAGCA 464
301 CCGTGGCCGAGCGGCTCATGAGCTGCTGCGGCTTCGTCGAGAGGAGAGGAGAGCA 360
465 GCGCGGCTCTGCGGCGGAGCGCTACCGCTGAGCAACCTGAGGAGAGGAGAGGAGAG 524
361 GCGCGGCTCTGCGGCGGAGCGCTACCGCTGAGCAACCTGAGGAGAGGAGAGGAGAG 420
525 CCGAGGCTGAGCGGAGCTGTAAGAGCGCTTCCAGCGGCGGAGAGGAGAGGAGAGAG 584
421 CCGTGGCTGAGCGGAGCTGTAAGAGCGCTTCCAGCGGAGAGGAGAGGAGAGGAGAG 480
585 GCGTTATGACTCGGAGAGCAACAGAGGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 644
481 GCGTTATGACTCGGAGAGCAACAGAGGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 540
645 TTGCGGCTGCGGAGCTGCTGAGCAACCTGTAAGAGCGCTGAGGAGAGGAGAGGAGAG 704
541 TACTGGCTGCGGAGCTGCTGAGCAACCTGTAAGAGCGGAGAGGAGAGGAGAGGAGAG 600
705 ATCCGGAGAGCTATCTGCTCAGAGAGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 764
601 ATCCGGAGAGAGCTATCTGCTCAGAGAGCTGTAAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 765 TGTGACACCTGTATGCTACAGCTGATTAATCCACCTGCTGTATACAGAGTGGTG 824
DB 661 TGTGAGATGCTCTTTCATCATGACAGGATTAACCCCTGCTGACAGGAGTGGTG 720
QY 825 ACTGTGGCGGTATACAGCTTCTTCTGACTTGTCTAGTTGGGCGGAGGAGTTGTGAACCA 884
DB 721 ACCATGCACTGTACAGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 885 GCCAAGGCTTACCTGCGGAGAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
DB 781 GCTCAGGCTTACCAAGACACGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 945 TTCTCTTCTTATGTTGGCTGCTGAGAGTGG 975
DB 841 TTCTCTTCTTACCGCGGCTGCTGAGAGTGG 871

RESULT 14
BC031186
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BC031186 1956 bp mRNA linear ROD 07-AUG-2002
Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
BC031186 GI:21411099
BC031186.1 GI:21411099
MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1956)
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) medpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 58 Row: 9 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analyses.
location/Qualifiers
1. 1956
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:37621 IMAGE:4989959"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_id="NCL CGAP_C024"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
176. 1573
/codon_start=1
/product="Unknown (protein for MGC:37621)"
/protein_id="AAH31186.1"
/db_xref="GI:21411100"
/db_xref="LocusID:212989"
/translation="MAISAAYRPLTSGKRFYKLYICDOVASLIPVSLFVYTLVVRNWSQYLCML
PALMCVAVGTHGRDNRRLYRPLIRAGLSAVLLRSVTAVERPDIHAYEA
GPMREERKREKFNLSINKYPCWWSNLAOARRGRIIDNSALKLELVEFR
GKGMFLHYDMISVPLVYQVVTALYSTFLACLIDRNFQVSLVDEKDYDLAYEK
DLVMDAERAPYATATVQLRPSFQSGSTFDITLAKEDMQRDLGDEPMEAPGD
FLDRLIPAGAGVAGPLGRRLSFLRKNSCVSEASTGSCSAVVEGADECSGD
PLDPLPEPBPAPGPPPLILPEPVPFSLVMPGPRGAPPLPSPIGEEBENT
A"

BASE COUNT	411 a	577 c	530 g	438 t	
ORIGIN					
Query Match	17.1%	Score 415	DB 10	Length 1956	
Best Local Similarity	67.0%	Pred. No. 6.9e-88			
Matches	589	Conservative	0	Mismatches 290	Indels 0
				Gaps 0	
97	GCCTGGCCATACATCACTTACACAGCCAAAGTGGCTTAATGCCGCTTAGGCTCTTCT	156	NSALKLLELVNFRSKCMLEHYMISIPYVTVYVAVSYFLACILGRFLDPA		
99	GCGCAGCATACCGCTACCTACACAGCCAGAGTGGGAATGCCGCTTGGTGGCTTCT	98	OGYKDHLDLCVPIFTLLOFFYAGWLAKAEOLINFGEDDDDEFENFLIDNRFQVM		
157	CCGCGCTGCTGCTGCTGCTGGGGGAGCAGCTATCAAGCTCTATATGGCGAGTCTTA	216	LAVDEYDLDLAMEKDYMDAAEAPAPYATATLLOQSPSGSPDATALAKSCCAG		
99	CCGACCTGCTGCTGCTGCTGGCGGAGACATCTACACCTCTGCGGAGAGCTGTAT	158	RUDGVDPLEVEHGFDFLORLLPAGAGSVGLPRRLSLLRKNSCSSEASTAKSCCAG		
217	TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	276	AADGGVGGCGDPLDPSLRPELEPPACPEPPAPIDGPTRPEFTTYSIPGRAPAP		
159	GTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	218	PMLPSPIGEEESPA"		
277	AACAGCTGATTTGAGAACTGACTGTATGCGACAGCTACATCCAGCTCATCCCA	336			
219	AGAACGCTACTTCCGAGAACTTGTATCTGCGACAGCAGCAGCAGCAGCAGCAGC	278			
337	TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	396			
279	TCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	338			
397	AGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	456			
339	TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	398			
457	ACGAGCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	516			
399	ACGATGAGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	458			
517	TGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	576			
459	TGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	518			
577	TGCAAGAGGCTTATATCTCCGAGCAACAAAGCTTGAGAACTGAGCTTACAC	636			
519	TGAGGCTGATTTATATACCGAGAGAGGCGCAAGAGTTCGAAACTTAAATTCG	578			
637	ACAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	696			
579	ACAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	638			
697	GAGGTGCAATCCGGAGCCCTTATCTGCTCCAGAGCTCTGAGAGAGTGAACCT	756			
639	AGGGCGGATCCCGACAAAGTGGCTTAAAGTTGCTATCAGAGAGCTGAATCT	698			
757	GTACTCAGTGGAGAACCTGTATGCTTACAGCTGATGATATATCCATGCTGTA	816			
699	GGAGCAATGTGGAGTCTGCTTACAGCTGATGATATATATCCATGCTGTA	758			
817	AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	876			
759	AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	818			
877	TGAACCCAGCCAAAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	936			
819	TGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	878			
937	TGCTGCAATGCTTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	975			
879	TGCTGCAATGCTTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	917			

RESULT 15	1957 bp	mRNA	linear	ROD 07-AUG-2002
BC019528				
LOCUS				
DEFINITION	Mus musculus, similar to hypothetical protein FLJ20132, clone			
ACCESSION	BC019528			
VERSION	BC019528.1			
KEYWORDS	GI:18044530			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 1957)			
JOURNAL	Strasberg, R.			
REMARK	Direct Submission			
COMMENT	Submitted (19-DEC-2001) National Institutes of Health, Mammalian			
	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-7590,			
	USA			
	NIH-MGC Project URL: http://mgc.ncl.nih.gov			
	Contact: MGC help desk			
	Email: cgaphs-rt@mail.nih.gov			
	Tissue Procurement: Jeffrey E. Green, M.D.			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)			
	DNA Sequencing by: Baylor College of Medicine Human Genome			
	Sequencing Center.			
	Center code: BCM-HGSC			
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/			
	Contact: amg@bcm.tmc.edu			
	Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Hale, S. M.,			
	Yoon, V. S., Kowis, C. R., Lawrence, S., Martin, R. G., Wuzny, D. M.,			
	Richards, S., Gibbs, R. A.			
	Clone distribution: MGC clone distribution information can be found			
	through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov			
	Series: IRAC Plate: 37 Row: h Column: 19			
	This clone was selected for full length sequencing because it			
	passed the following selection criteria: Similarity but not			
	identity to protein.			
FEATURES	location/Qualifiers			
source	1. 1957			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	/map="FVB/N"			
	/clone="MGC:28568 IMAGE:4208860"			
	/tissue_type="Colon, normal, 5 month old male mouse."			
	/clone_lib="MCI_CGAP.Co24"			
	/lab_host="DH10B"			
	/note="Vector: pCMV-Sport6"			
	189..1586			
	/codon_start=1			
	/product="Similar to hypothetical protein FLJ20132"			
	/protein_id="AAH19528.1"			
	/db_xref="GI:18044531"			
	/translation="MALSAAYFLLAEEGRKFEVLVYICQVSLIPSVFVGFYVT			
	LYVRMMNOYLCPDLDMALNCIVAGVHRDGRLLYRTIMRAGLSAVILHSVT			
	AVFRFTIDHVAEAPRMREERKFNENSSYKRYVPCVWFSSLLAQAAREGRID			
	NSALKLLELVNFRSKCMLEHYMDAAEAPAPYATATLLOQSPSGSPDATALAKSCCAG			
	OGYKDHLDLCVPIFTLLOFFYAGWLAKAEOLINFGEDDDDEFENFLIDNRFQVM			
	LAVDEYDLDLAMEKDYMDAAEAPAPYATATLLOQSPSGSPDATALAKSCCAG			
	RUDGVDPLEVEHGFDFLORLLPAGAGSVGLPRRLSLLRKNSCSSEASTAKSCCAG			
	AADGGVGGCGDPLDPSLRPELEPPACPEPPAPIDGPTRPEFTTYSIPGRAPAP			
	PMLPSPIGEEESPA"			

BASE COUNT	402 a	584 c	534 g	437 t	
ORIGIN					
Query Match	17.1%	Score 415	DB 10	Length 1957	
Best Local Similarity	67.0%	Pred. No. 6.9e-88			
Matches	589	Conservative	0	Mismatches 290	Indels 0
				Gaps 0	
97	GCCTGGCCATACATCACTTACACAGCCAAAGTGGCTTAATGCCGCTTAGGCTCTTCT	156	NSALKLLELVNFRSKCMLEHYMISIPYVTVYVAVSYFLACILGRFLDPA		

RESULT 15

BC019528

1957 bp. mRNA. linear. ROD 07-AUG-2002

LOCUS

Mus musculus, similar to hypothetical protein FLJ20132, clone

DEFINITION

MGC:28568 IMAGE:4208860, mRNA, complete cds.

ACCESSION

BC019528

VERSION

BC019528.1 GI:18044530

KEYWORDS

MGC.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE

1 (bases 1 to 1957)

JOURNAL

Strausberg, R.

REMARK

Direct Submission

COMMENT

Submitted (19-DEC-2001) National Institutes of Health, Mammalian

FEATURES

Gene Collection (MGC), Cancer Genomics Office, National Cancer

SOURCE

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

LOCATION/Qualifiers

USA

FEATURES

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

SOURCE

Contact: MGC help desk

FEATURES

Email: cgabs-remail.nih.gov

SOURCE

Tissue Procurement: Jeffrey E. Green, M.D.

FEATURES

CDNA Library Preparation: Life Technologies, Inc.

SOURCE

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

FEATURES

DNA sequencing by: Baylor College of Medicine Human Genome

SOURCE

Sequencing Center:

FEATURES

Center code: BCM-HGSC

SOURCE

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

FEATURES

Contact: angbcm.tmc.edu

SOURCE

Guaratine, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,

FEATURES

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

SOURCE

Richards, S., Gibbs, R.A.

FEATURES

Clone distribution: MGC clone distribution information can be found

SOURCE

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

FEATURES

Series: IRAC Plate: 37 Row: h Column: 19

SOURCE

This clone was selected for full length sequencing because it

FEATURES

passed the following selection criteria: similarity but not

SOURCE

identity to protein.

```
Db      52  GCGAGCGATGACCGTACCTACACAGCCAGAGTGGGAAATGCCCGCTCGGTGGCTTC 111
QY      157  CCGCGCTGCTGCTGCTGCGGGGACACATCTACAAAGCTGCTATATGGGAGTTCTTAA 216
Db      112  CCGAGCTGCTGCTGCTGCGGGGACATCTACAAAGCTGCTGCGGAGAGCTGTAT 171
QY      217  TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
Db      172  GTTTCCTGGGACTTACATGACCTAAGCGCCCTATCGCTTCTTACTGGCAGAGAGC 231
QY      277  AACAGCTGATGTTTGAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
Db      232  AGAAGCGCTACTTGGAGAGCTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 291
QY      337  TTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
Db      292  TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
QY      397  AGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
Db      352  TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
QY      457  ACGAGCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
Db      412  ACGATCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
QY      517  TCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
Db      472  TGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
QY      577  TGCAGCAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db      532  TCGAGGCTGATTTATGACCGAGAGGCGGAGAGAGTTGAGAACTTGATTCGTC 591
QY      637  ACAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db      592  ACAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
QY      697  GAGGTGATTCGGGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
Db      652  AAGGGGCGCATCCGCGACACAGTGCCTTAAGTTGCTACTAGAGAGCTGAATGTTC 711
QY      757  GTACTGATGCTGACACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Db      712  GGAGCAATGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
QY      817  AGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
Db      772  AGGTAGTCACTATCGAGTGTACGCTACTTCTTGGCTTGCATCGGCGCTCACTTC 831
QY      877  TGAACCGACCAAGGCTACCGTGGGCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 936
Db      832  TAGACCTGCGACAGGCTACAAAGACCAACCTGGAGCTATGCTATCCATCTTCACCT 891
QY      937  TCTGCAATCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
Db      892  TGTGCAATCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
```

Search completed: June 8, 2003, 18:10:36
Job time : 6129.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:51:03 ; Search time 3214.69 Seconds
(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-4
2429
Sequence: 1 caggagagtcacccagccta.....aaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_estr:*
9: gb_estl:*
10: gb_estl2:*
11: gb_estc3:*
12: gb_estc4:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	65.3	2453	11	BC015220 Homo sapi
2	640.4	26.4	666	13	BI480798 H2RE-043
3	582.4	24.0	592	13	BM663028 UI-E-C10-
4	577	23.8	773	10	BE410951 601303662
5	575	23.7	585	14	BM707948 UI-E-C11-
6	565.2	23.3	593	10	BE385296 601277572

7	564.6	23.2	963	14	BQ879880	BQ879880 AGENCOURT
8	541.2	22.3	1699	11	AK006549	AK006549 Mus muscu
9	528.8	21.8	735	13	BI756228	BI756228 603024265
10	525.8	21.6	537	14	BM691456	BM691456 UI-E-C11-
11	522	21.5	522	14	BM707649	BM707649 UI-E-C11-
12	507.4	20.9	930	14	BM685396	BM685396 UI-E-C10-
13	492.4	20.3	531	14	BQ436834	BQ436834 AGENCOURT
14	485.2	20.0	566	14	BM718338	BM718338 UI-E-BO1-
15	475.4	19.6	508	14	BM685122	BM685122 UI-E-EJ1-
16	475.4	19.6	508	14	BM932117	BM932117 UI-E-EJ1-
17	470.8	19.4	474	9	AI190190	AI190190 qd37c03.x
18	464.8	19.1	469	9	AA307119	AA307119 EST178031
19	444.6	18.3	601	13	BC951790	BC951790 MRI-CR073
20	443	18.2	481	13	BM662514	BM662514 UI-E-C11-
21	427.4	17.6	484	14	N31453	N31453 YX55608.f1
22	419	17.2	525	14	BQ345562	BQ345562 PM3-NT031
23	405.4	16.7	419	14	BM718146	BM718146 UI-E-EJ0-
24	391	16.1	766	10	BE275846	BE275846 601121656
25	378.8	15.6	1067	13	BM562042	BM562042 AGENCOURT
26	378.4	15.6	381	9	AI208843	AI208843 qg38g12.x
27	365.8	15.1	371	14	BM694530	BM694530 UI-E-C11-
28	360.6	14.8	603	9	AA205892	AA205892 zq4/a06.f
29	359.2	14.8	455	12	BG013943	BG013943 IL5-GN024
30	347.2	14.3	444	14	N33227	N33227 yy07602.s1
31	343.4	14.1	514	13	BE236901	BE236901 145906 MA
32	341.2	14.0	526	14	BI343182	BI343182 371242 MA
33	314	12.9	526	14	BM685501	BM685501 UI-E-C10-
34	310.4	12.8	793	13	BG961794	BG961794 602826439
35	307.4	12.7	599	13	BG951382	BG951382 MRI-CR073
36	299.4	12.3	317	14	N25339	N25339 YX55608.s1
37	293.4	12.1	305	9	AA318352	AA318352 EST20422
38	283.4	11.7	732	13	BG965806	BG965806 602829577
39	282.8	11.6	297	9	AA991578	AA991578 086010.8
40	271.6	11.2	394	13	BG950641	BG950641 MRI-CR073
41	246.2	10.1	271	9	AA317489	AA317489 EST19416
42	244.4	10.1	613	9	AL654606	AL654606 AL654606
43	243.4	10.0	422	10	AM594705	AM594705 hg63d05.x
44	239.8	9.9	705	14	BM719045	BM719045 UI-E-E01-
45	232	9.6	966	12	BF582222	BF582222 602094224

ALIGNMENTS

RESULT 1
LOCUS BC015220 2453 bp mRNA linear HTC 25-JUL-2002
DEFINITION Homo sapiens, clone IMAGE:3877806, mRNA.
ACCESSION BC015220
VERSION BC015220.1 GI:21955361
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2453)
Strauberg, R.
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DPF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 14 Row: b Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: incomplete processing.

FEATURES

source

Location/Qualifiers
1. 2453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3877806"
/issue_type="lung large cell carcinoma"
/clone_lib="NH_MGC_68"
/lab_host="DH10B"
/note="Vector: PCMV-Sport6"
BASE COUNT 633 a 711 c 606 g 503 t
ORIGIN

Query Match

Best Local Similarity 86.7% Pred. No. 3.6e-224; Length 2453;
Matches 1884; Conservative 0; Mismatches 5; Indels 284; Gaps 1;

```
OY 257 GCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGACTCTGATTGCGACAG 316
Db 557 GCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGACTCTGATTGCGACAG 616
OY 317 CTACATCCAGCTCATCCCATTTCTTCGTGCTGGCTTCTAGCTGACGCTGCTGAC 376
Db 617 CTACATCCAGCTCATCCCATTTCTTCGTGCTGGCTTCTAGCTGACGCTGCTGAC 676
OY 377 CCGCTGGTGAACAGATACGAGACCTGGCCGACCGCTCATGAGCCCTGCTGTC 436
Db 677 CCGCTGGTGAACAGATACGAGACCTGGCCGACCGCTCATGAGCCCTGCTGTC 736
OY 437 GGGCTTCGTCAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
Db 737 GGGCTTCGTCAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
OY 497 GCGCAACCTGGGCAACGCTGCTATCTGCGAGCGCTGACGCGAGCTCAAGAGCTT 556
Db 797 GCGCAACCTGGGCAACGCTGCTATCTGCGAGCGCTGACGCGAGCTCAAGAGCTT 856
OY 857 CCCCAGCGCCAGCACTGCTGCAAGCAAGAGGCTTTATGACTCCGGCAGAAACAGAGTT 916
Db 917 CCCCAGCGCCAGCACTGCTGCAAGCAAGAGGCTTTATGACTCCGGCAGAAACAGAGTT 976
OY 976 GGAAGAACTGAGCTTACACACACATGTTCTGGGTGCGCTGGGTGGGTGGGTGGGT 1036
Db 1037 GGAAGAACTGAGCTTACACACACATGTTCTGGGTGCGCTGGGTGGGTGGGTGGGT 1096
OY 1097 GATCAATGAAGGCGCTGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1156
Db 1157 GATCAATGAAGGCGCTGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1216
OY 1217 TATCCCACTGGTGTATACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1276
Db 1277 TATCCCACTGGTGTATACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1336
OY 1337 TATAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1396
Db 1397 TATAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1456
OY 1457 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1516
Db 1517 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1576
OY 1577 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1636
Db 1637 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1696
OY 1697 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1756
Db 1757 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1816
OY 1817 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1876
Db 1877 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1936
OY 1937 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 1996
Db 1997 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 2056
OY 2057 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 2116
Db 2117 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGGCCATGAGTGGACCT 2176
```

```
OY 977 CTTCTCCAGGCGCTGCTGGGTGAGAGCATGGCCAGAGGGGTATGCGCCAGCAGCTGCT 1036
Db 1270 ----- 1269
OY 1037 TGAGACGAGATGATGATGTCAGAGAAAGAGTCTACGGGTAGAAAGCAGCCAGCGT 1096
Db 1270 ----- 1269
OY 1097 GTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGAGAGATGCTGTAAAC 1156
Db 1270 ----- 1269
OY 1157 CGGAGCGGAGGTTGTGTGTCAGAGCAGCTATCAACCCCTTGGAGAGATGATGAT 1216
Db 1270 ----- 1269
OY 1270 ----- 1269
OY 1277 GAGATGACACAGGAGCTGCTCGATGAGGCGGACATGATGATGATGATGATGATGAT 1336
Db 1293 GAGATGACACAGGAGCTGCTCGATGAGGCGGACATGATGATGATGATGATGATGAT 1396
OY 1337 CAGCCCGCTTACAGAGCTGCTTCCGCGAGTTCGAGCTCTTATGAGGCTTCCAGC 1412
Db 1353 CAGCCCGCTTACAGAGCTGCTTCCGCGAGTTCGAGCTCTTATGAGGCTTCCAGC 1472
OY 1397 TTTCAATCATGAGCTGATCAAAAGAGAGATGAGTTCAGCCCATCAGAGAGAGAGAG 1456
Db 1413 TTTCAATCATGAGCTGATCAAAAGAGAGATGAGTTCAGCCCATCAGAGAGAGAGAG 1516
OY 1457 GATGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1532
Db 1473 GATGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1592
OY 1517 CCGAGGCGCAACTCAAGAGCAAACTGATGATGATGATGATGATGATGATGATGATGAT 1636
Db 1533 CCGAGGCGCAACTCAAGAGCAAACTGATGATGATGATGATGATGATGATGATGATGAT 1696
OY 1577 GCGCTGCGCAAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1756
Db 1593 GCGCTGCGCAAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1816
OY 1637 GCGTGAAGCTTAAAGCTGTGAGAGCGCTTCAAGTGTGAGAGAGAGAGAGAGAGAG 1696
Db 1653 GCGTGAAGCTTAAAGCTGTGAGAGCGCTTCAAGTGTGAGAGAGAGAGAGAGAGAG 1712
OY 1697 TACTACAGTCCCGCAGAGAGCGCCAGCCAGCTCCATGTTCTTCCCTTGAAGCA 1756
Db 1713 TACTACAGTCCCGCAGAGAGCGCCAGCCAGCTCCATGTTCTTCCCTTGAAGCA 1816
OY 1757 TCAAGCGCGTCAAAAGCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832
Db 1773 TCAAGCGCGTCAAAAGCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1896
OY 1817 ACTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1936
Db 1833 ACTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1996
OY 1877 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2056
Db 1893 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2116
OY 1937 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176
Db 1953 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2236
OY 1997 CACACTTACACTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
Db 2013 CACACTTACACTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2116
OY 2057 CATCTTACACTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2176
Db 2117 CATCTTACACTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2236
```

Db 2073 CATTCCTAACCTGCTTCTTAATGGGAGTCTTCCGACCCAGTCTCCTGCTGTGTAC 2132
QY 2117 ACCGACGAGACATGATCGATCGATCAGAGCATACAGCTGTCCACACTGAGAACGTCTCT 2176
Db 2133 ACCGACGAGACATGATCGATCGATCAGAGCATACAGCTGTCCACACTGAGAACATGTCTCT 2192
QY 2177 ACAACAGCCTGATCAATCAATGGTTAGCTTAATAGATATAAATCCAGACTACTTCAGCCTT 2236
Db 2193 ACAACAGCCTGATCAATCAATGGTTAGCTTAATAGATATAAATCCAGACTACTTCAGCCTT 2252
QY 2237 TAATGCTTTTATTTATATAAATCCTGTAAGCTAGTACGAAACATTTGAAACATTTACT 2296
Db 2253 TAATGCTTTTATTTATATAAATCCTGTAAGCTAGTACGAAACATTTGAAACATTTACT 2312
QY 2297 CAGACTGTGATTCAGAGTGGGAAACCTTACTTATCTGTAATCCAGACACACAC 2356
Db 2313 CAGACTGTGATTCAGAGTGGGAAACCTTACTTATCTGTAATCCAGACACACAC 2372
QY 2357 TTATGTTACTGCCCCAATATGAGTTTAAATATCAATATCTGTTAAAAAATAA 2416
Db 2373 TTATGTTACTGCCCCAATATGAGTTTAAATATCAATATCTGTTAAAAAATAA 2432
QY 2417 AAAAAAAAAA 2429
Db 2433 AAAAAAAAAA 2445

RESULT 2
BI480798 666 bp mRNA linear EST 28-FEB-2002
LOCUS HRP-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
DEFINITION 5' similar to vitelliform macular dystrophy (Best disease, mRNA
sequence.
VERSION BI480798
KEYWORDS EST.
SOURCE BI480798.1 GI:18998607
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 666)
REFERENCE
AUTHORS Mammals; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Yuan, Y., Macnee, S.P., Hughes, B. and Swaroop, A.
Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100
genes
JOURNAL
COMMENT Unpublished (2001)
CONTACT: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kelllogg Eye Center, University of Michigan
540 KEE, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu
PCR Primers
FORWARD: M13/PUC-Reverse - ccagtcacagctgttaaacg
BACKWARD: M13/PUC-Forward - agcgatacaattccacacag
Seq primer: M13/PUC-Reverse.

FEATURES
source
1. .666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Retinal Pigment Epithelium (2)"
/tissue_type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/note="Organ: Retina; Vector: pSport1"
BASE COUNT 133 a 224 c 171 g 136 t 2 others
ORIGIN
Query Match 26.4%; Score 640.4; DB 13; Length 666;
Best Local Similarity 99.4%; Pred. No. 6.7e-85;
Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 10 CCACAGCTAGTGCAGACCTTCTGTGGATCATGAGCCACCTGGAACCCACCTG 69
Db 12 CCACAGCTAGTGCAGACCTTCTGTGGATCATGAGCCACCTGGAACCCACCTG 71
QY 70 ACCCAAGCCACCTGCTGCAGACCCACTGCTGGCCATGACATCACTTACAGAGCAAG 129
Db 72 ACCCAAGCCACCTGCTGCAGACCCACTGCTGGCCATGACATCACTTACAGAGCAAG 131
QY 130 TGGCTATAGCCGCTTAGGCTCTTCTCCGCTGCTGTGCTGCGGGGAGCATCT 189
Db 132 TGGCTATAGCCGCTTAGGCTCTTCTCCGCTGCTGTGCTGCGGGGAGCATCT 191
QY 190 ACAAGCTGTATATGCGAGTTCTTAATCTTCTGCTGTCTACTACATCACTGCTTA 249
Db 192 ACAAGCTGTATATGCGAGTTCTTAATCTTCTGCTGTCTACTACATCACTGCTTA 251
QY 250 TTTTATAGCTGGCCCTCAGGAAAGAACACAGCTGATTTTGAAGAACTGACTTAT 309
Db 252 TTTTATAGCTGGCCCTCAGGAAAGAACACAGCTGATTTTGAAGAACTGACTTAT 311
QY 310 GCGACAGCTATACATCAGCTATCCCATTTCTTCTGCTGTGCGCTTATAGTACGCTG 369
Db 312 GCGACAGCTATACATCAGCTATCCCATTTCTTCTGCTGTGCGCTTATAGTACGCTG 371
QY 370 TCGTGACCCGCTGTGTGAGAACCATGAGAACCTGCGTGGCCGACGCTCATGAGCC 429
Db 372 TCGTGACCCGCTGTGTGAGAACCATGAGAACCTGCGTGGCCGACGCTCATGAGCC 431
QY 430 TGTGTGCGGCTGTGTGAGAACCATGAGAACCTGCGTGGCCGACGCTCATGAGCC 489
Db 432 TGTGTGCGGCTGTGTGAGAACCATGAGAACCTGCGTGGCCGACGCTCATGAGCC 491
QY 490 TCCGCTAGCCACCTGCTGCGCAAGCTGCTCATCTGCGAGAGGTAGAGACGCACTTACA 549
Db 492 TCCGCTAGCCACCTGCTGCGCAAGCTGCTCATCTGCGAGAGGTAGAGACGCACTTACA 551
QY 550 AGCGCTTCCGCGCCGCGCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
Db 552 AGCGCTTCCGCGCCGCGCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 610
QY 610 ACCAGTTGAGAAATGAGCTTACACACACATGTTGGTGGCTGCTGCTGCTGCTGCTG 665
Db 611 ACCAGTTGAGAAATGAGCTTACACACACATGTTGGTGGCTGCTGCTGCTGCTGCTG 666

RESULT 3
BM663028/c 592 bp mRNA linear EST 27-FEB-2002
LOCUS UI-E-C10-aad-h-10-0-UI-s1 UI-E-C10 Homo sapiens cDNA clone
DEFINITION UI-E-C10-aad-h-10-0-UI 3', mRNA sequence.
VERSION BM663028
KEYWORDS EST.
SOURCE BM663028.1 GI:18968017
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 592)
REFERENCE
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
COMMENT Genome Res. 6 (9), 791-806 (1996)
MEDLINE
CONTACT: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Ecksstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msocres@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L16M287 row: 0 column: 18
High quality sequence stop: 593.

FEATURES

source

1..593
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3618655"
/clone_lib="NIH-MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; site:1: XhoI; site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 176 a 177 c 132 g 108 t
ORIGIN

Query Match 23.3%; Score 565.2; DB 10; Length 593;
Best Local Similarity 99.3%; Pred. No. 7.9e-74;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1471 CATCATGGGGGCTTCCTAGGCGCTGCAGTCCCATGATACCATCTCTCCAGGGCAATC 1530
1 CATCATGGGGGCTTCCTAGGCGCTGCAGTCCCATGATACCATCTCTCCAGGGCAATC 60
1531 AAGGACCAACTACTGTGGCCCAAGAGGAATCCCTCTCCAGAGGGCTGCCC -AAA 1589
61 AAGGACCAACTACTGTGGCCCAAGAGGAATCCCTCTCCAGAGGGCTGCCC -AAA 120
1590 ACCCAAGGCGCAACACAGAGCTTAGGGCCAGGAAGACACAGAGGCTGGAAGCTTA 1649
121 ACCCAAGGCGCAACACAGAGCTTAGGGCCAGGAAGACACAGAGGCTGGAAGCTTA 180
1650 AGGCTGTGAGCGCTTCAAGCTGCGCCACAGTATCAAGAGGCGGCTACTACAGTCCC 1709
181 AGGCTGTGAGCGCTTCAAGCTGCGCCACAGTATCAAGAGGCGGCTACTACAGTCCC 240
1710 CACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCGCTGAACATCAGCGCGTCAA 1769
241 CACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCGCTGAACATCAGCGCGTCAA 300
1770 AGCTTCACTGTACACAGGATAGACACCAAGCAAAAGCTTAAGACTGTGATCTG 1829
301 AGCTTCACTGTACACAGGATAGACACCAAGCAAAAGCTTAAGACTGTGATCTG 360
1830 GGGCGCAAGAAAGTTGTAATGCTCTCAAGAGCGATGGGGCTTATGAGACACCA 1889
361 GGGCGCAAGAAAGTTGTAATGCTCTCAAGAGCGATGGGGCTTATGAGACACCA 420
1890 AAGTATCTCAAGTGAAGAGAAACTGTGAGTTAAGTACGAGGATATGAGAGATCC 1949
421 AAGTATCTCAAGTGAAGAGAAACTGTGAGTTAAGTACGAGGATATGAGAGATCC 480
1950 CCGAATAATCACTCAAGAACTTTGGAACAATACCAACCAACATACACTACACTCA 2009
481 CCGAATAATCACTCAAGAACTTTGGAACAATACCAACCAACATACACTACACTCA 540
2010 AAGATACATGATCTTATTTGGGCTTGGAAAAACAGAGATG 2051
541 AAGATACATGATCTTATTTGGGCTTGGAAAAACAGGCTG 582

RESULT 7

BQ879880

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BQ879880 963 bp mRNA linear EST 16-ATG-2002
AGENCOURT 8241531 lupskl_dorsal_root_ganglion Homo sapiens CDNA
clone IMAGE:6180559 5', mRNA sequence.
BQ879880 BQ879880.1 GI:22271888
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 963)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L16M13563 row: 1 column: 08
High quality sequence stop: 623.

FEATURES

source

1..963
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6 (Life Technologies); Site:1: NotI; Site:2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCG-3' and 5'-GACATGTTCTGATGATCGGACGGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
BASE COUNT 205 a 289 c 265 g 202 t 2 others
ORIGIN

Query Match 23.2%; Score 564.6; DB 14; Length 963;
Best Local Similarity 77.3%; Pred. No. 7.7e-74;
Matches 795; Conservative 0; Mismatches 31; Indels 203; Gaps 1;

445 TCGAAGCAAGAGGACAGAGCGGCTGCTGCGGCGCACGCTACCGGCAACC 504
26 TCGAAGCAAGAGGACAGAGCGGCTGCTGCGGCGCACGCTACCGGCAACC 85
505 TGGGCAAGTGTCTCATCTCGGCGGCTGAGACCGGAGCTCAACGCTTCCCGAGCG 564
86 TGGGCAAGTGTCTCATCTCGGCGGCTGAGACCGGAGCTCAACGCTTCCCGAGCG 145
565 CCGAGCACTGTGTCAAGAGGCTTATGACTCCGCGAGAACACAGCTTGGAGAAAC 624
146 CCGAGCACTGTGTCAAGAGGCTTATGACTCCGCGAGAACACAGCTTGGAGAAAC 205
625 TGAAGCTACACACAACTTTGTTGGTGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 684
206 TGAAGCTACACACAACTTTGTTGGTGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 265
685 AGGCTGTGAGTTCGAATCCGGAGCCCTTACCTGCTCCAGAGGCTGTAAGAGA 744

Db	266	AGGGGCGCTTGGAGTGCATCCGGGACCCCTATTCCTGCTCCAGACCCCTGCTGAACGAGA	325	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636
Qy	745	TGAACACCTTTCCTACTGACAGTGGACACCTGATGCTACAGACTGATGATATCCAC	804		
Db	326	TGAACACCTTTCCTACTGACAGTGGACACCTGATGCTACAGACTGATGATATCCAC	385		
Qy	805	TGGTGTATACAGAGTGGTACATGTGGCGGTGACAGCTTCTCTCACTTGTCTAGTTG	864	TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Db	386	TGGTGTATACAGAGTGGTACATGTGGCGGTGACAGCTTCTCTCACTTGTCTAGTTG	445	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Itou, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
Qy	865	GGCGGAGTTTCTGAACCCAGCCAGGCTACCTGGCCATGAGCTGGACCTGCTTGTGC	924		
Db	446	GGCGGAGTTTCTGAACCCAGCCAGGCTACCTGGCCATGAGCTGGACCTGCTTGTGC	505		
Qy	925	CCGCTTTCACGCTTCTGCAAGTCTTCTTATGATGGCTGCTGAAGTGGGCTCTCA	984		
Db	506	CCGCTTTCACGCTTCTGCAAGTCTTCTTATGATGGCTGCTGAAGTGGGCTCTCA	551		
Qy	985	GGGCGCTGCTGGGCTGGAGCATGGCCAGAGGGGTCTATGGCCACAGACTCTTGACGCA	1044	TITLE	Shibata, K., Itou, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Db	552	GGGCGCTGCTGGGCTGGAGCATGGCCAGAGGGGTCTATGGCCACAGACTCTTGACGCA	551	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Kono, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itou, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Itoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
Qy	1045	GGATGAGTGTGAGAAAGAGTCTCAGCGGTAGAAAGACAGCAGCGTGGTGGCGCA	1104		
Db	552	GGATGAGTGTGAGAAAGAGTCTCAGCGGTAGAAAGACAGCAGCGTGGTGGCGCA	551		
Qy	1105	CACCTGTATCCAGCTACTCGGAGGCTGAGGAGAGAGATGCTTGAACCGGAGGC	1164	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itou, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K., Saito, T., Okazaki, Y., Gojodori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Glass, C., Kling, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustlincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-Oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilmberg, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S., and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
Qy	1225	GACCAACTGAGATTTGACAGAGAAATTTGAGAGTCTCTTGGCTGTGGATGATGCA	1284		
Db	603	GACCAACTGAGATTTGACAGAGAAATTTGAGAGTCTCTTGGCTGTGGATGATGCA	662		
Qy	1285	CCAGAGACCTGCTCGAGAGGAGCCGACATGATGATTAAGCCCGAGCCACAGCCGCC	1344	TITLE	Shinagawa, A., Shibata, K., Yoshino, M., Itou, M., Ishii, Y.,
Db	663	CCAGAGACCTGCTCGAGAGGAGCCGACATGATGATTAAGCCCGAGCCACAGCCGCC	722	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Arakawa, T., Balderelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, Y., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itou, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shitaki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, K., I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y. Direct submission Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
Qy	1465	CGCTGGCAT 1473	842		
Db	843	CTCCCGCCT 851			
RESULT 8	AK006549	1699 bp	LINEAR	TITLE	Shinagawa, A., Shibata, K., Yoshino, M., Itou, M., Ishii, Y.,
LOCUS	AK006549	1699 bp	LINEAR	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Arakawa, T., Balderelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, Y., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itou, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shitaki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, K., I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y. Direct submission Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
ACCESSION	AK006549	1699 bp	LINEAR		
VERSION	AK006549	1699 bp	LINEAR		
KEYWORDS	AK006549	1699 bp	LINEAR		
SOURCE	AK006549	1699 bp	LINEAR		
ORGANISM	AK006549	1699 bp	LINEAR		
REFERENCE	AK006549	1699 bp	LINEAR		
AUTHORS	AK006549	1699 bp	LINEAR		
TITLE	AK006549	1699 bp	LINEAR		

FEATURES
source

BASE COUNT
ORIGIN

Query Match	22.38;	Score 541.2;	DB 11;	Length 1699;
Best Local Similarity	66.78;	Pred. No. 1.6e-70;		
Matches 983; Conservative	0;	Mismatches 248;	Indels 243;	Gaps 5

OY 273 GAGCAAGACGAGAGTGGTTGAGAAACTGACTCTGTATTGGACAGCTACATCCAGCTCATC 332
 Db 1 GATCAGCAGCGTGTGTTTGAAGAGGTGGCTCTGTACTGCGACAGCTACATTCAGCTCATC 60
 OY 333 CCCATTTTCCTTGTCGTGGGCTTCTACGTGAGCGTGGTCGACACCGGTGGTGAACAG 392
 Db 61 CCTATATCTCTGTTCTGGGTTTCTATGTATTACATTGGTGGTGAAGCCGCTGGTGGACCA 120
 OY 393 TACGAGAACCTTCGCGTGGCCGACCGCCTCATGAGCCTGGTGTGCGGCTTGCATCAAGC 452
 Db 121 TACGGAAGACTTCGCGTGGCCGACCGCCTCATGATCAGAGTGTACGTTGTTGAGGAGG 180
 OY 453 AAGGACGAGCAAGGGCCGCTGCGGGGGGACGCGTCATCCGCTACGCCAACCTGGGCAAC 512
 Db 181 AAGATGAGGAAGAGCCGTTTCTGCGGGGACGCGTCATCCGCTACGCGCATCCTGGGCCAA 240
 OY 513 GTGCTCATCTGCGCAGCGCTGACGACCGCAGCTTACAAGCGCTTCCCGACGCCAGCAC 572
 Db 241 GTGCTCATCTCGCGCAGCATCAGCACCTCGGCTCTACAAGCGCTTCCCACTCTTCCACAC 300
 OY 573 CTGGTGCAGCAGGCTTATGATCTCCGCGACAACAAGCAGTTGGAGAAACTGACCTA 632
 Db 301 CTGGTGCAGCAGGCTTATGATCAACCATGAGGGGACAAATAGCAGTTGCAAAAGTTGGGCTA 360
 OY 633 CCACACAACATGTTCTGGGTCCCTGGGGTGGTTTGGCAACGCTCAATAAGAGCGGG 692
 Db 361 CCACA -AACACATTCTGGGTGCCCTGGGTGGTTTGGCAACCTTCAATTAAGGCGCTAT 419
 OY 693 CTTGGAGCGTAATCCGGAGCCCTATCTGCTCCAGAGCGCTGAAAGAGATGAACACC 752
 Db 420 CTTGGAGCGTAATCCGGAGCACCGCTGCTGCTCCAGAGCGTGATTAAGAGCTGTGACT 479
 OY 753 TTGGGTACTAGTGTGGACACCTGTATGCTTACAGACTGAGATTAGTATTCACCTGTGTAT 812
 Db 812

[illegible]

VERSION B1756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.lnl.gov>
 Plate: LMNL1487 row: a column: 18
 High quality sequence stop: 577.
 Location/Qualifiers
 1. 735

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194649"
 /clone_1lb="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb. Insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

BASE COUNT 135 a 229 c 214 g 157 t
 ORIGIN

Query Match 21.8%; Score 528.8; DB 13; Length 735;
 Best Local Similarity 92.6%; Pred. No. 1.6e-68;
 Matches 613; Conservative 0; Mismatches 37; Indels 12; Gaps 5;

257 GCTGGCCCTCAGGAGAAACAAGCTGATGTTGAGAACTGACTCTGATTTGGACAG 316
 Db 59 GCTGGCCCTCAGGAGAAACAAGCTGATGTTGAGAACTGACTCTGATTTGGACAG 118
 317 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGAGCTTCTACGACGCTGGTCTGAC 376
 Db 119 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGAGCTTCTACGACGCTGGTCTGAC 178
 377 CCGCTGTGGAACAGTACGAGACCTGCGGGCCGCGCCGCTCATAGCGTGGTGTTC 436
 Db 179 CCGCTGTGGAACAGTACGAGACCTGCGGGCCGCGCCGCTCATAGCGTGGTGTTC 238
 437 GGGCTTCTCGAAGGCAAGAGGAGGAGGCGGCTGCTGCGGCGACGCTCATCGGCA 496
 Db 239 GGGCTTCTCGAAGGCAAGAGGAGGAGGCGGCTGCTGCGGCGACGCTCATCGGCA 298
 497 CGCCAACTGGGCAACGCTGCTATCTCTGCGCAGCGCTACAGCGAGTCTACAGCGCTT 556
 Db 289 CGCCAACTGGGCAACGCTGCTATCTCTGCGCAGCGCTACAGCGAGTCTACAGCGCTT 358
 557 CCCAGGCGCCAGCACTGTGTGCAAGAGGCTTTATGACTCCGGCAAGAACACACAGCT 616
 Db 359 CCCAGGCGCCAGCACTGTGTGCAAGAGGCTTTATGACTCCGGCAAGAACACACAGCT 418
 617 GGAAGAACTGAGCTTACACACAAACATGTTCTGGGTGCGCTGGGTGTTGCCAAGCT 676
 Db 419 GGAAGAACTGAGCTTACACACAAACATGTTCTGGGTGCGCTGGGTGTTGCCAAGCT 478
 677 GTCAATGAAAGCGTGGGTGAGGTGCAATCCGGGACCTTATCTGCTCAGAGGCTGCT 736

Db 479 GTCAATGAAAGCGTGGGTGAGGTGCAATCCGGGACCTTATCTGCTCAGAGGCTGCT 538
 737 GAACGAGTGAACACCTT-GGGTACTAGTGTGAGACACTGTATGCTT-ACAGCTGATTT 794
 Db 539 GACGAGTGAACACCTTGGGCTGCTGCTGAGACACTGTATGCTTACAGCTGATTT 598
 795 AGTATCCCACTGGGTATATACAC-----AGTGTGTAAGTGTGGGCTGATAGCTTCTT 847
 Db 599 AGTATCCCACTGGGTATATATCCCGGCTGGGTGATACCTGTGGGCTGCTGCTTCTT 658
 848 CCTGACTGTCTAGTTGGG--CGGAGTTTGTGAA-CCGACCAAGGCTTACCTCGGCCA 904
 Db 659 CCTGACTGTCTAGTTGGGCGGCGGCTTGTGAACCCCGCAGGAGGCTTCCCTGGGCT 718
 905 TG 906
 Db 719 TG 720

RESULT 10
 BM691456 537 bp mRNA linear EST 28-FEB-2002
 LOCUS
 DEFINITION
 UI-E-CII-abe-f-10-0-UI-r1 UI-E-CII Homo sapiens CDNA clone
 ACCESSION
 BM691456
 VERSION
 BM691456.1 GI:19004714
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 537)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE
 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seg primer: M13 Reverse.

FEATURES
 source

1. 537
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-CII-abe-f-10-0-UI"
 /clone_1lb="UI-E-CII"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)."
 /note="Organ: eye; Vector: pRT3-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-CII is a normalized CDNA library containing the
 following tissue(s): RPE and choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand CDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded CDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pRT3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand CDNA contains a library tag
 sequence that is located between the Not I site and the

Query Match 20.3%; Score 492.4; DB 14; Length 930;
Best Local Similarity 97.4%; Pred. No. 3.2e-63;
Matches 532; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 1189 CATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCAACGATGTCGACAGAA 1248
DB 99 CATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCAACGATGTCGACAGAA 138
QY 1249 TTTCGAGGTGTCCCTGTGGCTGTGATGATGACACAGACCTGCTCGATGAGCC 1308
DB 159 TTTCGAGGTGTCCCTGTGGCTGTGATGATGACACAGACCTGCTCGATGAGCC 218
QY 1309 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
DB 219 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 278
QY 1369 CCGTCGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACCAAGAGAGATGA 1428
DB 279 CCGTCGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACCAAGAGAGATGA 338
QY 1429 GTTCCAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 1488
DB 339 GTTCCAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 398
QY 1489 AGGCTGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACCAAGAGAGATGA 1548
DB 399 AGGCTGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACCAAGAGAGATGA 458
QY 1549 GCCCAAG 1608
DB 459 GCCCAAG 518
QY 1609 GAACCTTAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 1667
DB 519 GAACCTTAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 578
QY 1668 AGCTGGGCGGCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1725
DB 579 AGCTGGGCGGCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
QY 1726 CCCCCC 1731
DB 639 ACCCCC 644

RESULT 14
BM718338 566 bp mRNA linear EST 01-MAR-2002
LOCUS UI-E-E01-a1a-1-23-0-UI-1 UI-E-E01 Homo sapiens cDNA clone
DEFINITION BM718338
ACCESSION BM718338
VERSION 1
KEYWORDS Normalization and subtraction: two approaches to facilitate gene
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 566)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source
1. 566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-E01-a1a-1-23-0-UI"
/clone_1b="UI-E-E01"
/issue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site: 1: EcoR I; Site: 2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(GT)₁₈ tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."

BASE COUNT 107 a 182 c 145 g 132 t
ORIGIN

Query Match 20.0%; Score 485.2; DB 14; Length 566;
Best Local Similarity 99.4%; Pred. No. 4.5e-62;
Matches 487; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 CTATGTCGAG 77
DB 1 CTATGTCGAG 60
QY 78 CCACCTGTCGAG 137
DB 61 CCACCTGTCGAG 120
QY 138 GCCCGCTTAGGCTCTCTCCGCTGCTGTGCTGGGCGGAGAGATCTACAAGCTG 197
DB 121 GCCCGCTTAGGCTCTCTCCGCTGCTGTGCTGGGCGGAGAGATCTACAAGCTG 180
QY 198 CTATGTCGAG 257
DB 181 CTATGTCGAG 240
QY 258 CTGCGCTTCACGAG 317
DB 241 CTGCGCTTCACGAG 300
QY 318 TACATTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 377
DB 301 TACATTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 360
QY 378 CGCTGTCGAG 437
DB 361 CGCTGTCGAG 420
QY 438 GCGTTCGTCGAG 497
DB 421 GCGTTCGTCGAG 480
QY 498 GCCAAGCTGG 507
DB 481 GCCAAGCTGG 490

RESULT	15
LOCUS	Bm685122/c
DEFINITION	Bm685122 503 bp mRNA linear EST 27-FEB-2002
VERSION	UI-E-EJ1-a}1-1-09-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
ACCESSION	BM685122
KEYWORDS	BM685122.1 GI:18995018
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Euarchonta; Primates; Carnivora; Homnidae; Homo. 1 (bases 1 to 503) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)
JOURNAL MEDLINE	97044477
COMMENT	Contact: Soares, MB

Email: msources@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq. Primer: M13 Forward
pOLIA-yes.

FEATURES

Source

```

BASE COUNT
ORIGIN
96 a 107 c 123 g 177 t

/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="UI-E-EJ1-aj1-1-09-0-U1"
/clone_1lb="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foreal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab="Morgan-DH10B (Life Technologies) (T1 phage resistant)"
/note="Host: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: Ecor I; Site.2: Not I;
UI-E-EJ1 is a subtracted cDNA library constructed
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tags for this library are: fetal eyes, AATACATGCA
; lens, CGATTGCGA; eye anterior segment, AATCGCATG;
optic nerve, GTCG; RPE and Choroid, ACCGA; Retina foreal and
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_1lb=UI-E-EJ1
TAG_Tissue=RPE and Choroid
TAG_Seq=ACCGA"

```

Query Match	19.6%	Score 475.4;	DB 14;	length 503;
Best Local Similarity	99.8%;	Pred. No. 1.3e-60;		
Matches 476; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

1750 AGAACCATCAGCGCGCTCAAAGCTTCACAGTGTACACAGGCATAGACACCAAGACAAG 1809

[illegible]

Search completed: June 9, 2003, 03:30:06
Job time : 3218.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:10:38 ; Search time 291.723 Seconds

(without alignments)
10693.617 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaa 2229

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1247.2	56.0	1263	9 US-09-746-783-3	Sequence 3, Appl1
2	350.4	15.7	1326	9 US-10-071-766-10	Sequence 10, Appl1
3	215.4	9.7	1350	10 US-09-768-826-16	Sequence 16, Appl1
4	189.2	8.5	1292	9 US-10-198-846-11070	Sequence 11070, A
5	175	7.9	1198	10 US-09-880-107-2174	Sequence 11346, A
6	116	5.2	853	9 US-10-198-846-11346	Sequence 11346, A
7	104	4.7	462	9 US-10-198-846-2561	Sequence 2561, Ap
8	102	4.6	615	9 US-10-198-846-7843	Sequence 7843, Ap
9	98	4.4	507	9 US-09-764-872-195	Sequence 195, Ap
10	64	2.9	751	9 US-10-198-846-1862	Sequence 1862, Ap
11	49.2	2.2	1263	9 US-09-746-783-3	Sequence 3, Appl1
12	49.2	2.2	1326	9 US-10-071-766-10	Sequence 10, Appl1
13	47	2.1	2076	9 US-09-986-480-116	Sequence 116, Appl
14	46.6	2.1	846	10 US-09-349-015-25	Sequence 25, Appl
15	46.6	2.1	8195	9 US-10-156-239-7	Sequence 7, Appl1
16	46.6	2.1	8195	9 US-10-199-485-7	Sequence 7, Appl1
17	46.6	2.1	8195	9 US-09-795-693-7	Sequence 7, Appl1
18	45.8	2.1	2338	10 US-09-925-302-90	Sequence 90, Appl
19	45.4	2.0	506	9 US-09-918-995-7423	Sequence 7423, Ap

20	45.4	2.0	667	9 US-09-813-153-77	Sequence 77, Appl
21	44.8	2.0	2341	9 US-09-991-053-9	Sequence 9, Appl1
22	44.8	2.0	2341	9 US-09-957-187-9	Sequence 9, Appl1
23	44.8	2.0	2607	9 US-09-991-053-11	Sequence 11, Appl
24	44.8	2.0	2607	9 US-09-957-187-11	Sequence 11, Appl
25	44.4	2.0	4909	9 US-10-042-141-25	Sequence 25, Appl
26	44.4	2.0	4909	10 US-09-726-643-25	Sequence 25, Appl
27	44.2	2.0	827	10 US-09-925-301-488	Sequence 488, App
28	44.2	2.0	912	10 US-09-764-853-142	Sequence 142, App
29	44.2	2.0	913	10 US-09-764-853-148	Sequence 148, App
30	44.2	2.0	400	10 US-09-920-300A-1253	Sequence 1253, App
31	44	2.0	400	12 US-10-033-528-1253	Sequence 1253, Ap
32	44	2.0	1342	9 US-09-879-389B-1	Sequence 1, Appl1
33	44	2.0	15832	9 US-10-239-676-106	Sequence 106, App
34	43.8	2.0	189	9 US-10-060-036-2380	Sequence 2380, Ap
35	43.8	2.0	931	9 US-09-991-036-1575	Sequence 1575, Ap
36	43.2	1.9	442	10 US-09-960-352-9688	Sequence 9688, Ap
37	43.2	1.9	1356	9 US-09-991-053-31	Sequence 31, Appl
38	43.2	1.9	1356	9 US-09-957-187-31	Sequence 31, Appl
39	43.2	1.9	2185	9 US-09-992-598-228	Sequence 228, App
40	43.2	1.9	2185	9 US-09-982-293A-228	Sequence 228, App
41	43.2	1.9	2185	9 US-09-982-735-228	Sequence 228, App
42	43.2	1.9	2185	9 US-09-990-444-228	Sequence 228, App
43	43.2	1.9	2185	9 US-09-989-730-228	Sequence 228, App
44	43.2	1.9	2185	9 US-09-990-436-228	Sequence 228, App
45	43.2	1.9	2185	9 US-09-991-181-228	Sequence 228, App

ALIGNMENTS

RESULT 1
US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Racle, Lisa A.
Trecay, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive

CITY: Cambridge
STATE: MA

COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 742-7400

TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO. 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs

RESULT 3

US-09-768-826-16
 ; Sequence 16, Application US/09768826
 ; Patent No. US20020012966A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shi et al.
 ; TITLE OF INVENTION: 18 human secreted proteins
 ; FILE REFERENCE: PF512P1
 ; CURRENT APPLICATION NUMBER: US/09/768,826
 ; CURRENT FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: PCT/US00/22350
 ; PRIOR FILING DATE: 2000-08-15
 ; PRIOR APPLICATION NUMBER: 60/148,759
 ; PRIOR FILING DATE: 1999-08-16
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 1350
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1135)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1148)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1166)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1174)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1181)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1209)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1229)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1266)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1285)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1287)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1290)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1295)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1305)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1324)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1339)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1341)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1343)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE

LOCATION: (1345)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (1348)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-768-826-16

Query Match 9.7%; Score 215.4; DR 10; Length 1350;
 Best Local Similarity 61.9%; Pred. No. 3.8e-53;
 Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3;

QY	603	GAACACAGCAGTGGAGAACTGACCTTACCACACATGTTGGGTCGGGCGG	662
DB	23	GGACGGGTGGGTTGAGAGCTGAATCCGACTTCAACAGTACCTGCTGCTC	82
QY	663	TGGTTCACCACTGTCATGAAGCGTGGCTGGAGTGCATCCGGACCTTAC	722
DB	83	TGGTTCACCACTGTCATGAAGCGTGGCTGGAGTGCATCCGGACCTTAC	142
QY	723	CTCCAGAGCTGCTGAACAGATGAACACCTTGGCTGATGATGATGATG	782
DB	143	CTCTGTCTACTTTTGGAGAGCTGAACAGTACCAAGCCAGATGATGATG	202
QY	783	TACGACTGATTAATATCCACTGGTATACACAGTGTGACTGTGGGCTG	842
DB	203	TATGACTGATCAGATCCCTCCCTCTACACCAAGTGTGATGATGATG	262
QY	843	TTCTTCCTGACTTGTCTAGTTGGGGGAGTTTGAACCCAG-----	885
DB	263	TTCTTGGCCCTCTCTGTTGGCCGCGAGTTGTGAGCCAGAGGGGCTC	322
QY	886	-----CCAAGGCTACCCCTGGCCATGAG-----CTGACCTC	917
DB	323	CCTCAGAGCTTCTGAAGCCAGGCGAGGACGCCAGCCCTGGAGACCC	382
QY	918	GTTGGCCGCTTACGTTCTCTGATGTTGCTGCTGCTGCTGCTGCTG	977
DB	383	TACGCTCTCACCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	442
QY	978	GAGCAGCTCATCAACCCCTTGTGAGAGATGATGATTTTGGAGCCAC	1037
DB	443	GACAGATCATCAACCCCTTGTGAGAGATGATGATTTTGGAGCCAC	502
QY	1038	GACAGATTTTGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1097
DB	503	GACCGCAACTGCGAGTGTCTCTCTATCCGTGAGCAAAATGACAACT	562
QY	1098	ATGAGCGCGACATGATGATGATGATGATGATGATGATGATGATG	1154
DB	563	GCTGAGAGAGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	622
QY	1155	TCCGCCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1214
DB	623	GCGGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	682
QY	1215	GAGAGATGAGTTCACG 1233	
DB	683	GACCTGAGCAGAGCTGC 701	

RESULT 4

US-10-198-846-11070
 ; Sequence 11070, Application US/10198846
 ; Publication No. US2003009974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steilmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-049

```

; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070
```

```

Query Match      8.5%; Score 189.2; DB 9; Length 1292;
Best Local Similarity 62.4%; Pred. No. 2.3e-45;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
```

```

QY 349 TGGGCTTCTAGCTAGACGCTGCTGACCCGCTGCTGACACCAAGTACGAGAACCCTGCCGT 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 TAGGCTTTATGATCTCTGTGTAGTGAACCGATGCTGGAACCAAGTTGTGAATTTGCCCT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 409 GGGCCGACCGCCCTCATGAGCCTGGTGTGCGGCTTCGTGAGAGGACGAGGAGGAGCC 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GGGCAGACAGGCTATGTTCTCTCATCTCTAGACATGTTTACAGGACGACGACGAGGCG 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 469 GCGTGTCTGGCGGCGACGCTCATCCGCTACGCGCAACCTGGGACAGCTGTCTCTCGGCA 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GCGTGTCTGAGAGGACGCTCATGCGCTACGCTCATCTCCTGCTCATCTTTCCT 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 529 GCGTACAGCAGCGAGCTATACAGGCGTTCGCCAGCGCCGACCTGCTGTCAGACAGCT 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CGGTGACACATGCTGTGTACAAAAGATTTCCACAAAGAGACACAGTGTGGAACGAGTT 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 TTTAGTACTCCGCGAGAACACAGAGAGTGGAGAACTGAGCCTACACACAACTGTCT 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 TTTATGACAAAGATGAAAGAAATTTATTAACACGCTCAAGTCTCCCATCTGAAATTT 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 649 GGGGCGCCCTGGGTGTGTTGCCAACCTGTCAAAGAGAGGCTGTGGAGGTGGAATCC 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 GGGTTCATTCATCTGTGTTGAAATCTTCAACTAAAGCCGGAAGGAGTGAATCA 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 GGGACCTTATCTCTCCAGAGCCTGCTGAACGATGAACACCTGCTGACTCAGTGTG 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 GAGACAGTGTGTGATCGCATCATGTGATGACTGAAGATTCGATACCGCTCTTGGTCA 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 769 GACACCTGTATGCTCTACGACTGATTAATCCCATCTGTGTATACACAGGTG 822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 GCCTCTTATTCGGTTATGACTGGTGTGCGTCCGCTGTTTACACCCAGGTAG 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 5
US-09-880-107-2174/C
; Sequence 2174, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles In Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2174
```

```

; LENGTH: 1198
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174
```

```

Query Match      7.9%; Score 175; DB 10; Length 1198;
Best Local Similarity 100.0%; Pred. No. 3.6e-41;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2027 CAGCCTTATGCTTTTATTCATTAACCTGTGAAGCTAGACTGAACCTTGAACA 2086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1198 CAGCCTTATGCTTTTATTCATTAACCTGTGAAGCTAGACTGAACCTTGAACA 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2087 TTTACACAGACTGTGATTCAGAGTGGGAAACCTTAGTTCTATCTGAATCAAGACAG 2146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1138 TTTACACAGACTGTGATTCAGAGTGGGAAACCTTAGTTCTATCTGAATCAAGACAG 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2147 CCACACCTTAGTATCTGCCCAACTAATGATTAATTAATCAATACCTGTT 2201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1078 CCACACCTTAGTATCTGCCCAACTAATGATTAATTAATCAATACCTGTT 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 6
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhilber, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11346
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
; LOCATION: 43, 50, 51, 54, 57, 853
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346
```

```

Query Match      5.2%; Score 116; DB 9; Length 853;
Best Local Similarity 68.2%; Pred. No. 1e-23;
Matches 161; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```

QY 350 GGGCTTCTAGCTAGACGCTGCTGACCCGCTGTGGAACCAAGTACGAGAACCCTGCCGTG 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 GGGGTTTATGTTACTCTGTGTGTAACCGATGGTGAACCAAGTTGTGAATTTCCCTG 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 410 GCGGACCGCCCTCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 GCGACACAGGCTTAATGTTCTCATCTCTAGAGTGTTCACGAAACCGACGAGCGGCG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 470 GCTGTGCGGGGACAGCTATCCGTACGCAACCTGGGCAACGTCATCTCTGCGGAG 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 CCGTGTGAGAGAGAGCGATGATGCGGTACGTAATCTACCTCCCTGCTCATCTTTCGCTC 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 530 GGTGAGACCGGACGCTCAAGAGCGCTTCCGAGCGCCGACGCTGTGTGAAGCAG 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 GGTGAGACGCTGTGTACAAAAGATTTCCACAAATGAGACGACGCTGTGAAGCAG 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:29:42 ; Search time 100.098 Seconds

(without alignments)
7441.920 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429
Sequence: 1 caggagagtcaccacagccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/prodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/1/lna/ECTUS.COMB.seq:*
6: /cgn2_6/prodata/1/lna/backfills1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	84.6	3.5	5590	US-09-050-159-129	Sequence 129, App
C 2	84.6	3.5	9365	US-09-608-285A-8	Sequence 8, Appli
C 3	84.6	3.5	9365	US-09-350-836B-8	Sequence 8, Appli
C 4	84.6	3.5	9365	US-09-370-265-8	Sequence 8, Appli
C 5	84.6	3.5	14747	US-09-608-285A-42	Sequence 42, Appli
C 6	84.6	3.5	15977	US-09-608-285A-59	Sequence 59, Appli
C 7	84.6	3.5	50000	US-09-146-053-3	Sequence 3, Appli
C 8	84.2	3.5	11531	US-08-068-945A-1	Sequence 1, Appli
C 9	84.2	3.5	11531	US-08-442-806-1	Sequence 1, Appli
C 10	83.2	3.4	36741	US-09-301-665-3	Sequence 3, Appli
C 11	83	3.4	1701	US-09-078-294-9	Sequence 9, Appli
C 12	83	3.4	2115	US-08-395-800A-7	Sequence 7, Appli
C 13	83	3.4	6769	US-08-480-784-20	Sequence 20, Appli
C 14	83	3.4	6769	US-08-483-553-20	Sequence 20, Appli
C 15	83	3.4	6769	US-08-467-002-20	Sequence 20, Appli
C 16	83	3.4	6769	US-08-483-554B-20	Sequence 20, Appli
C 17	83	3.4	6769	US-08-488-011B-20	Sequence 20, Appli
C 18	83	3.4	6769	US-08-850-727-20	Sequence 20, Appli
C 19	83	3.4	6769	PCT-US95-10202-20	Sequence 20, Appli
C 20	83	3.4	6769	PCT-US95-10203-20	Sequence 20, Appli
C 21	83	3.4	6769	PCT-US95-10220-20	Sequence 20, Appli
C 22	83	3.4	14636	US-09-173-914-6	Sequence 6, Appli
C 23	83	3.4	20674	US-09-641-658-651	Sequence 651, App
C 24	83	3.4	70000	US-09-851-896-3	Sequence 3, Appli
C 25	83	3.4	112132	US-09-741-150-3	Sequence 3, Appli
C 26	82.6	3.4	282	US-08-133-629-8	Sequence 8, Appli
C 27	82.6	3.4	1613	US-08-812-204-1	Sequence 1, Appli

C 28	82.6	3.4	11725	US-08-756-506-1	Sequence 1, Appli
C 29	82.6	3.4	38564	US-09-734-673-3	Sequence 3, Appli
C 30	82	3.4	59065	US-09-813-817-3	Sequence 3, Appli
C 31	82	3.4	59065	US-09-978-197-3	Sequence 3, Appli
C 32	82	3.4	99500	US-09-798-096-10	Sequence 10, Appli
C 33	82	3.4	111282	US-09-754-250-3	Sequence 3, Appli
C 34	81.4	3.4	4421	US-08-257-963B-9	Sequence 9, Appli
C 35	81.4	3.4	4421	US-08-367-841A-6	Sequence 6, Appli
C 36	81.4	3.4	4421	US-08-520-373D-6	Sequence 6, Appli
C 37	81.4	3.4	4421	US-09-978-197-3	Sequence 3, Appli
C 38	81.4	3.4	5581	US-09-973-544-1	Sequence 1, Appli
C 39	81.4	3.4	5789	US-09-242-948-3	Sequence 3, Appli
C 40	81.4	3.4	7705	US-08-687-080-115	Sequence 115, App
C 41	81.4	3.4	12394	US-09-593-995-10	Sequence 10, Appli
C 42	81.4	3.4	20598	US-09-743-588-3	Sequence 3, Appli
C 43	81.4	3.4	36159	US-08-965-048-5	Sequence 5, Appli
C 44	81.4	3.4	45716	US-08-965-048-5	Sequence 5, Appli
C 45	81.4	3.4	45989	US-08-965-048-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT 1
US-09-050-159-129/c
Sequence 129, Application US/09050159A
Patent No. 6197505
GENERAL INFORMATION:
APPLICANT: NO. 6197505berg, Ielf T
APPLICANT: Andersson, Maria K
APPLICANT: Ilnstrom, Per H
TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
FILE REFERENCE: 1248/10042
CURRENT APPLICATION NUMBER: US/09/050,159A
CURRENT FILING DATE: 1998-03-27
EARLIER APPLICATION NUMBER: 60/042,930
EARLIER FILING DATE: 1987-04-03
NUMBER OF SEQ ID NOS: 133
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 129
LENGTH: 5590
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
US-09-050-159-129
Query Match 3.5%; Score 84.6; DB 4; Length 5590;
Best Local Similarity 90.9%; Pred. No. 5.9e-11;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1079 AGAAGCAGCCAGCGGTGTGGCGCACACCTGTATCCAGCTACTCGGAGGCTAGGC 1138
DB 1926 AAAAAATTAGCAGCGGTGTGGCGGTTCCTGTATCCAGCTACTCGGAGGCTAGGC 1867
QY 1139 AGAAGATCCCTTGAACCCGCGGAGCGGAGGCTGTGTG 1177
DB 1866 AGGAGATCCCTTGAACCCGCGGAGCGAGAGATGTGTG 1828
RESULT 2
US-09-608-285A-8/c
Sequence 8, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
```

```

; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; US-09-608-285A-8

Query Match          3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 7.3e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCGCGGTGGTGGCGACACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTAGCCGCGGTGGTGGCGACACCTGTAATCCAGCTACTCGGAGGCTGAGGC 7539
QY 1139 AGAGATCGCTTGAACCCGCGGAGCGGAGGCTTGTGTG 1177
DB 7538 AGRAGATCGCTTGAACCCGAGAGGAGGAGGTGCACTG 7500

RESULT 3
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
```

```

; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1) .. (288)
; NAME/KEY: exon
; LOCATION: (1281) .. (1580)
; NAME/KEY: exon
; LOCATION: (1820) .. (1855)
; NAME/KEY: exon
; LOCATION: (2467) .. (2555)
; NAME/KEY: exon
; LOCATION: (2863) .. (2942)
; NAME/KEY: exon
; LOCATION: (3889) .. (3950)
; NAME/KEY: exon
; LOCATION: (4894) .. (4995)
; NAME/KEY: exon
; LOCATION: (5847) .. (5987)
; NAME/KEY: exon
; LOCATION: (6966) .. (7138)
; NAME/KEY: exon
; LOCATION: (8556) .. (9365)
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; US-09-350-836B-8

Query Match          3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 7.3e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCGCGGTGGTGGCGACACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTAGCCGCGGTGGTGGCGACACCTGTAATCCAGCTACTCGGAGGCTGAGGC 7539
QY 1139 AGAGATCGCTTGAACCCGCGGAGCGGAGGCTTGTGTG 1177
DB 7538 AGRAGATCGCTTGAACCCGAGAGGAGGAGGTGCACTG 7500

RESULT 4
US-09-370-265-8/c
; Sequence 8, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
```

EARLIER FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: 09/244,444
EARLIER FILING DATE: 1999-02-04
EARLIER APPLICATION NUMBER: 09/122,449
EARLIER FILING DATE: 1998-07-24
EARLIER APPLICATION NUMBER: 09/118,205
EARLIER FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8

LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (72)..(287)
FEATURE:
NAME/KEY: CDS
LOCATION: (1280)..(1579)
FEATURE:
NAME/KEY: CDS
LOCATION: (1819)..(1854)
FEATURE:
NAME/KEY: CDS
LOCATION: (2466)..(2555)
FEATURE:
NAME/KEY: CDS
LOCATION: (2863)..(2940)
FEATURE:
NAME/KEY: CDS
LOCATION: (3887)..(3952)
FEATURE:
NAME/KEY: CDS
LOCATION: (4896)..(4994)
FEATURE:
NAME/KEY: CDS
LOCATION: (5846)..(5986)
FEATURE:
NAME/KEY: CDS
LOCATION: (6965)..(7138)
FEATURE:
NAME/KEY: CDS
LOCATION: (8556)..(8639)
FEATURE:
NAME/KEY: misc feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-370-265-8

Query Match 3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 7.3e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCAGCGGTGGTGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTAGCCRGCGGTGGTGGCGGCGCTGTATCCAGCTACTCGGAGGCTGAGGC 7539
QY 1139 AGGAGATCGCTTGACCCCGGAGGCGGAGGCTGTGTG 1177
DB 7538 AGRAGATCGCTTGAAACCAAGAGAGGAGGAGGTGACGTG 7500

RESULT 5
US-09-608-285A-42/c
Sequence 42, Application US/09608285A
Patent No. 6335013

GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
TITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

Query Match 3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 8.9e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCAGCGGTGGTGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 10958 AAAAATTAGCCRGCGGTGGTGGCGGCGCTGTATCCAGCTACTCGGAGGCTGAGGC 10899
QY 1139 AGGAGATCGCTTGACCCCGGAGGCGGAGGCTGTGTG 1177
DB 10898 AGRAGATCGCTTGAAACCAAGAGAGGAGGTGACGTG 10860

RESULT 6
US-09-608-285A-59/c
Sequence 59, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
TITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 59
LENGTH: 15977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CD39-L4/L66 Gene Sequence

NAME/KEY: CDS
LOCATION: (245)..(461)
NAME/KEY: CDS
LOCATION: (1454)..(1533)
NAME/KEY: CDS
LOCATION: (2734)..(2877)
NAME/KEY: CDS
LOCATION: (4364)..(4439)
NAME/KEY: CDS
LOCATION: (4679)..(4714)
NAME/KEY: CDS
LOCATION: (5326)..(5414)
NAME/KEY: CDS
LOCATION: (5723)..(5802)
NAME/KEY: CDS
LOCATION: (6751)..(6812)
NAME/KEY: CDS
LOCATION: (7758)..(7859)
NAME/KEY: CDS
LOCATION: (8712)..(8852)
NAME/KEY: CDS
LOCATION: (9831)..(9887)
NAME/KEY: CDS
LOCATION: (11613)..(11728)
NAME/KEY: CDS
LOCATION: (13146)..(13691)
NAME/KEY: CDS
LOCATION: (15702)..(15839)
NAME/KEY: misc_feature
LOCATION: (14871)
OTHER INFORMATION: n = a or c or g or t
US-09-608-2854-59

Query Match 3.5%; Score 84.6; DB 4; Length 15977;
Best Local Similarity 84.8%; Fred. No. 9.2e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
Db 1079 AGAAGCAGCAGCGTGTGGCGCACACTGTATCCAGCTACTGGAGCGTAGGC 1138
12188 AAAAATAGCGCGGCGTGTGGCGCACACTGTATCCAGCTACTGGAGCGTAGGC 12129
QY 1139 AGGAGATCGCTTGAACCGGAGCGGAGCGAGGTGTGTG 1177
12128 AGGAGATCGCTTGAACCGGAGCGGAGCGAGGTGTGTG 12090

RESULT 7
US-09-146-053-3/c
Sequence 3, Application US/09146053A
Patent No. 639349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C.
TITLE OF INVENTION: Human Aminopeptidase P Gene
FILE REFERENCE: MGI103
CURRENT APPLICATION NUMBER: US/09/146,053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057,854
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 3
LENGTH: 50000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-3

Query Match 3.5%; Score 84.6; DB 4; Length 50000;
Best Local Similarity 90.9%; Fred. No. 1.5e-10;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 1079 AGAAGCAGCAGCGTGTGGCGCACACTGTATCCAGCTACTGGAGCGTAGGC 1138
8689 AAAAATAGCGCGGCGTGTGGCGCACACTGTATCCAGCTACTGGAGCGTAGGC 8630
QY 1139 AGGAGATCGCTTGAACCGGAGCGGAGCGAGGTGTGTG 1177
1139 AGGAGATCGCTTGAACCGGAGCGGAGCGAGGTGTGTG 1177
Db 8629 AGGAGATCGCTTGAACCGGAGCGGAGGTGTGTG 8591

RESULT 8
US-08-068-945A-1
Sequence 1, Application US/08068945A
Patent No. 5616483
GENERAL INFORMATION:
APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Eneback, Sven
APPLICANT: Hansson, Lennart
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
APPLICANT: Tornell, Jan
TITLE OF INVENTION: New DNA Sequences
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,945A
FILING DATE: 27-MAY-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9202088-2

FILING DATE: 03-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9300902-5
 FILING DATE: 19-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Sterner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)819-8783
 TELEFAX: (212)354-8113
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11531 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: Mammary gland
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
 OTHER INFORMATION: /EC_number= 3.1.1.1
 OTHER INFORMATION: /product= "Bile Salt-Stimulated Lipase"
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..1640
 FEATURE:
 NAME/KEY: TATA_signal
 LOCATION: 1611..1617
 FEATURE:
 NAME/KEY: exon
 LOCATION: 1641..1727
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4071..4221
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4307..4429
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4707..4904
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6193..6323
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6501..6608
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6751..6868
 FEATURE:
 NAME/KEY: exon
 LOCATION: 8335..8521
 FEATURE:
 NAME/KEY: exon
 LOCATION: 8719..8922
 FEATURE:
 NAME/KEY: exon
 LOCATION: 10124..10321
 FEATURE:
 NAME/KEY: exon
 LOCATION: 10650..11490
 FEATURE:

NAME/KEY: 3'UTR
 LOCATION: 11491..11531
 US-08-068-945A-1
 Query Match
 Best Local Similarity 91.8%; Score 84.2; DB 1; Length 11531;
 Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1081 AAAGCAGCCAGCGCTGTGTCGCCACACCTGTATCCCACTACTCGGAGGCTGAGGAG 1140
 DB 5361 AATATGACGAGCGGTGTGTCGCCCTTGCCCTGTATCCCACTACTCGGAGGCTGAGGAG 5420
 QY 1141 GAGATTCGCTTGACCCGGAGCGGAGGCTGTGTG 1177
 DB 5421 GAGATTCGCTTGACTGACGAGCGGAGGCTGTGCGTG 5457
 RESULT 9
 US-08-442-806-1
 Sequence 1, Application US/08442806
 Patent No. 5716817
 GENERAL INFORMATION:
 APPLICANT: Bjursell, Gunnar
 APPLICANT: Carlsson, Peter
 APPLICANT: Enerback, Sven
 APPLICANT: Hansson, Lennart
 APPLICANT: Lidberg, Ulf
 APPLICANT: Nilsson, Jeanette
 APPLICANT: Tornell, Jan
 TITLE OF INVENTION: Genomic DNA Sequences
 TITLE OF INVENTION: Encoding Human BSSL/CEL
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/442,806
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/068,945
 FILING DATE: 27-MAY-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9201809-2
 FILING DATE: 11-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9201826-6
 FILING DATE: 12-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9202088-2
 FILING DATE: 03-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9300902-5
 FILING DATE: 19-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Sterner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)819-8783
 TELEFAX: (212)354-8113
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:

LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
FEATURE:
NAME/KEY: mat.peptide
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /EC_number=3.1.1.1
FEATURE:
NAME/KEY: /Product= "Ble salt-stimulated lipase"
LOCATION: 5'UTR
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1611..1617
FEATURE:
NAME/KEY: exon
LOCATION: 1641..1727
FEATURE:
NAME/KEY: exon
LOCATION: 4071..4221
FEATURE:
NAME/KEY: exon
LOCATION: 4307..4429
FEATURE:
NAME/KEY: exon
LOCATION: 4707..4904
FEATURE:
NAME/KEY: exon
LOCATION: 6193..6323
FEATURE:
NAME/KEY: exon
LOCATION: 6501..6608
FEATURE:
NAME/KEY: exon
LOCATION: 6751..6868
FEATURE:
NAME/KEY: exon
LOCATION: 8335..8521
FEATURE:
NAME/KEY: exon
LOCATION: 8719..8922
FEATURE:
NAME/KEY: exon
LOCATION: 10124..10321
FEATURE:
NAME/KEY: exon
LOCATION: 10650..11490
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 11491..11531

Query Match 3.5%; Score 84.2; DB 1; Length 11531;
Best Local Similarity 91.8%; Pred. No. 1e-10;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1081 AAAGCAGCAGCGCTGTGGCGCACACCTGTAAATCCAGCTACTCGGAGAGCTGAGGCAG 1140
DB 5361 AAATTAGCAGCGCTGTGGCGCTGTGCTGTAAATCCAGCTACTCGGAGAGCTGAGGCAG 5420
OY 1141 GAGAAATCGCTTGAACCCGGAGGCGGAGGCTTGTGCTG 1177

DB 5421 GAGAAATCGCTTGAACCTCAGGAGGCGGAGGCTTGCCTG 5457

RESULT 10
US-09-301-665-3/c
Sequence 3, Application US/09301665
Patent No. 6207876
GENERAL INFORMATION:
APPLICANT: KELLEMS, RODNEY E.
APPLICANT: DATA, SURJIT K.
APPLICANT: BLACKBURN, MICHAEL R.
TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
FILE REFERENCE: METHODS FOR THE USE THEREOF
CURRENT APPLICATION NUMBER: US/09/301,665
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: 60/083,408
EARLIER FILING DATE: 1998-04-29
EARLIER APPLICATION NUMBER: 60/083,370
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 36741
TYPE: DNA
ORGANISM: Homo sapiens
US-09-301-665-3

Query Match 3.4%; Score 83.2; DB 4; Length 36741;
Best Local Similarity 80.8%; Pred. No. 2.9e-10;
Matches 97; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 1081 AAAGCAGCAGCGCTGTGGCGCACACCTGTAAATCCAGCTACTCGGAGAGCTGAGGCAG 1140
DB 1531 AAATTAGCAGCGCTGTGGCGCTGTGCTGTAAATCCAGCTACTCGGAGAGCTGAGGCAG 1472
OY 1141 GAGAAATCGCTTGAACCCGGAGGCGGAGGCTTGTGCTGACAGCAGCTATCAACCCCTT 1200
DB 1471 GAGAAATCGCTTGAACCTCAGGAGGCTTGTGCTGACAGCAGCTATCAACCTATTCATT 1412

RESULT 11
US-09-078-294-9/c
Sequence 9, Application US/09078294
Patent No. 6265211
GENERAL INFORMATION:
APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Sart, Desiree
APPLICANT: Cancilla, Michael R.
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
FILE REFERENCE: Davies Col
CURRENT APPLICATION NUMBER: US/09/078,294
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 1701
TYPE: DNA
ORGANISM: BAC-F2 contig 5
US-09-078-294-9

Query Match 3.4%; Score 83; DB 4; Length 1701;
Best Local Similarity 89.9%; Pred. No. 8.6e-11;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1079 AGAAGCAGCAGCGCTGTGGCGCACACCTGTAAATCCAGCTACTCGGAGAGCTGAGGC 1138
DB 1274 AAATTAGCAGCGCGCTGTGGCGCTGTGCTGTAAATCCAGCTACTCGGAGAGCTGAGGC 1215
OY 1139 AGAAGCAGCAGCGCTGTGGCGCGGAGGCGGAGGCTTGTGCTG 1177
DB 1214 AGAAGCAGCAGCTTGAACCCGGAGGCGGAGGCTTGCAGTG 1176

```
RESULT 12
US-08-395-800A-7
; Sequence 7, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROUCQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELE: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..1092
; US-08-395-800A-7

Query Match
Best Local Similarity 89.9%; Pred. No. 9.4e-11; Length 2115;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCAGCGGTGTGGCGCACACCTGTAATCCAGCTACTGGAGGCTGAGGC 1138
DB 1375 AAAAATATAGCCAGCGGTGTGGCGCACACTGTATATCCAGCTACTGGAGGCTGAGGC 1434
QY 1139 AGGAGATCGCTTGAACCCGGGAGCGGAGGCTGTGTG 1177
DB 1435 AAGAGATCCTTGAACCCAGGAGCGGAGGTTGCAGTG 1473

RESULT 13
US-08-480-784-20
; Sequence 20, Application US/08480784
; Patent No. 5693473
; GENERAL INFORMATION:
; APPLICANT: SKOLNICK, MARK H.
; APPLICANT: GOLDSBART, DAVID E.
; APPLICANT: MIKI, YOSHIO
; APPLICANT: SWENSON, JEFF
; APPLICANT: KAMB, ALEXANDER
; APPLICANT: HARTSHMAN, KEITH D.
```

```
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavligian, Sean V.
APPLICANT: Wiegman, Roger W.
APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,784
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6769 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-480-784-20

Query Match
Best Local Similarity 82.6%; Pred. No. 1.6e-10; Length 6769;
Matches 95; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCAGCGGTGTGGCGCACACCTGTAATCCAGCTACTGGAGGCTGAGGC 1138
DB 2396 AAAAATATAGCCAGCGGTGTGGCGCACACTGTATATCCAGCTACTGGAGGCTGAGGC 2455
QY 1139 AGGAGATCGCTTGAACCCGGGAGCGGAGGCTGTGTGCGCAGAGCAGCTCATCA 1193
DB 2456 AGGAGATCCTTGAACCCGGGAGGCTGTGTGCGGTAACCCAGATCGACCA 2510

RESULT 14
US-08-483-553-20
```

```

Db      2396 AAAAAATACGAGGTGTGGCATACATCTATATCCAGCTACTCGGAGGCTGAGGC 2455
QY      1139 AGGAGATCGCTTGAACCCGGAGCGAGGTTGTGGTGGCAGAGCACTCATCA 1193
        ||||||| ||||||| ||||||| ||||||| |||  | || |
Db      2456 AGGAGATCATCTTGAACCCGGAGGTGGAGGTTGGCGGTGAACCGAGATCGACCA 2510

RESULT 15
US-08-487-002-20
/ Sequence 20, Application US/08487002
/ Patent No. 5710001
/ GENERAL INFORMATION:
/ APPLICANT: Shattuck-Eidens, Donna M.
/ APPLICANT: Shardt, Jacques
/ APPLICANT: Eml, Mitsuru
/ APPLICANT: Nakamura, Yusuke
/ APPLICANT: Durocher, Francine
/ TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
/ NUMBER OF SEQUENCES: 85
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
/ STREET: 1201 New York Avenue, N.W., Suite 1000
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/487,002
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/409,305
/ FILING DATE: 24-MAR-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/348,824
/ FILING DATE: 29-NOV-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/308,104
/ FILING DATE: 16-SEP-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/300,266
/ FILING DATE: 02-SEP-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/289,221
/ FILING DATE: 12-AUG-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ihnen, Jeffrey L.
/ REGISTRATION NUMBER: 28,957
/ REFERENCE/DOCKET NUMBER: 24884-109347
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-962-4810
/ TELEFAX: 202-962-8300
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6769 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHEetical: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ US-08-487-002-20

Query Match      3.4%; Score 83; DB 1; Length 6769;
Best Local Similarity 82.6%; Pred. No. 1.6e-10;

```

	Matches	95;	Conservative	0;	Mismatches	20;	Indels	0;	Gaps	0;
QY	1079	ACAAAGCAGCCAGCGCTGCTGGCGCACACCTGTATCCAGCTACTCGGAGGCTGAGGC	1138							
Db	2396	AAAAATTAGCCAGGTGTGTGGCACATTAATCCAGCTACTCGGAGGCTGAGGC	2455							
QY	1139	AGGAGATCGCTGACCCGCGGAGGAGTGTGTGGCAGAGCACTCATCA	1193							
Db	2456	AGGAGATCACTGTAAACCGGAGGTGTGGTGAACCCAGATCGCACCA	2510							

Search completed: June 8, 2003, 18:16:31
 Job time : 102.098 secs